Habit Formation as an Explanation for the Equity Premium in Germany

A thesis presented to the faculty of economics,

University of Mannheim

in partial fulfillment of the requirements for the

degree of Diplomvolkswirt

Chair of Economic Theory: Prof. Felix Kübler, Ph.D.

Manuel Wirtz

Mannheim, November 3rd, 2005

Abstract

I discuss if the consumption-based asset pricing model is able to explain the equity premium in Germany. I use data for the period from 1960 to 1994. The result is that this is not the case. So there is an equitiy premium puzzle in Germany. The habit formation model of Campbell and Cochrane (Campbell and Cochrane 1999) is calibrated and discussed for German data.

Deklaration

I declare that the thesis is written independently and without the use of other supplies than mentioned. The following programs were used: Excel, GAUSS, MATLAB and LaTeX.

Contents

1 Introduction

The term "Equity Premium Puzzle" goes back to Mehra and Prescott (Mehra and Prescott 1985). The consumption-based asset pricing model is unable to explain the excess return of stocks over the short term riskfree interest rate. If the consumption-based model is applied to postwar equity returns and riskfree interest rates one needs an unrealistically high value of risk aversion in the CRRA utility function[1].
Excess returns of stocks over the riskfree interest rate are predictable especially in the long run. So there would be no justification for an equity premium in the long run. The variables which predict excess returns are correlated with the business cycle or predict them (Fama and French 1989). Equity risk premia seem to be higher at business cycle troughs than they are at peaks. Price-dividend ratios move procyclically, but this movement cannot be explained by variation in expected excess returns (Campbell and Shiller 1988a,b). Because this movement cannot be explained by changes of expected dividends or interest rates there must be large countercyclical variation in expected excess returns. Estimates of conditional variances change over time (Bollerslev et al. 1992). Conditional variances do not move one-for-one with estimates of conditional mean returns. So the slope of the conditional mean-variance frontier, which is a measure for the price of risk, changes through time with a business cycle pattern. Equivalently, the Sharpe ratio of mean variance efficient returns also changes with a business cycle pattern. In subsequent research there are different approaches to that problem.
Campbell and Cochrane (Campbell and Cochrane 1999) show a solution using external habit formation. Habit formation comes from psychology. If agents get used to a certain level of consumption they take it for granted and fear constellations that would leave them with a level of consumption close to habit or below habit. As a result people care more about changes of consumption than about the level. Or in the language of Campbell and Cochrane: repetition of a stimulus diminishes the perception. So in macroeconomics habit formation can explain why recessions are so feared even though their effects on output may be relatively small after some years of growth.
Consumption growth and dividend growth are modeled as independently identically distributed (i.i.d.) lognormal processes. Habit is always below consumption and utility is a function of consumption and habit $u(C - X)$. In other models like Sundaresan (Sundaresan 1989), Ferson (Ferson and Constantinides 1991), Heaton (Heaton

[1] CRRA is an acronym for constant relative risk aversion

1995) and Chapman (Chapman 1997) habit can fall below consumption. Habit formation is external and enters the model by a heteroskedastic first order autoregressive process. Habit moves slowly in response to consumption and adapts nonlinearly to the history of consumption. The real riskfree rate is constant by construction. A constant riskfree rate is consistent with a linear production technology and makes it possible to close the model without the addition of an explicit production sector. The model predicts many puzzles that face the standard power utility consumption-based model, including the equity premium and riskfree rate puzzles and the low unconditional correlation of consumption growth and stock returns. The model is consistent with an even sharper long-run equity premium puzzle that results from mean-reversion in stock prices, together with low long-run consumption volatility.
In the next chapter I present the consumption-based asset pricing model. In chapter three I derive the algebra and especially the Hansen-Jagannathan bound. Chapters three and four present some basic knowledge and are adapted from the book "Asset Pricing" (Cochrane 2001). The shortcomings of the consumption-based model are described in chapter four. In chapter five I present German data and find out if the standard consumption-based model works with German data. In chapter six the habit formation model of Campbell and Cochrane is presented. It is a short presentation of the paper of Campbell and Cochrane (Campbell and Cochrane 1999). In chapter seven I apply the habit formation model to German data and check if it works. Chapter eight is a critical review of the habit formation model. Chapter nine concludes the thesis. Since it is appropriate to show how the model is constructed, I have written several appendices which lead the reader through the algebra. All equations used in the working paper version as well as in the version published in the Journal of Political Economy are derived.

2 The Consumption-Based Model

The utility of a respresentative agent is a function of current and future values of consumption:

$$U(c_t, c_{t+1}) = u(c_t) + E_t[\beta u(c_{t+1})]. \tag{2.1}$$

The utility function is assumed to be of the CRRA[1] form:

$$u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma} \tag{2.2}$$

The limit as $\gamma \to 1$ is

$$u(c_t) = \ln(c_t). \tag{2.3}$$

A representative agent maximizes:

$$\max_{\xi} \quad U(c_t, c_{t+1}) = u(c_t) + E_t[\beta u(c_{t+1})] \quad s.t. \tag{2.4}$$

$$c_t = e_t - p_t\xi$$

$$c_{t+1} = e_{t+1} + x_{t+1}\xi$$

Where c_t is its consumption in period t and β is a subjective discount factor. Its income in period t is e_t and the amount of assets it chooses to buy is ξ. Substituting the constraints into the objective function and setting the derivative with respect to ξ to zero, we obtain the first-order condition for an optimal consumption and portfolio choice,

$$p_t u'(c_t) = E_t[\beta u'(c_{t+1}) x_{t+1}] \tag{2.5}$$

[1]CRRA means constant relative risk aversion.

$$p_t = E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}x_{t+1}\right]. \tag{2.6}$$

Define the *stochastic discount factor* m_{t+1} as

$$m_{t+1} \equiv \beta\frac{u'(c_{t+1})}{u'(c_t)}. \tag{2.7}$$

The pricing formula can be expressed as

$$p_t = E_t(m_{t+1}x_{t+1}). \tag{2.8}$$

If the asset payoff is nonstochastic, then

$$p_t = E_t(m_{t+1}x_{t+1}) \quad \Leftrightarrow \quad p_t = \frac{1}{R^f}x_{t+1}. \tag{2.9}$$

3 Asset Pricing

3.1 The Sharpe Ratio

The basic pricing equation is:

$$p = E(mx) \tag{3.1}$$

Using the definition of the covariance $cov(m, x) = E(mx) - E(m)E(x)$ we can write:

$$p = E(m)E(x) + cov(m, x) \tag{3.2}$$

Substituting the riskfree rate:

$$p = \frac{E(x)}{R^f} + cov(m, x) \tag{3.3}$$

Discounted gross returns of an arbitrary asset i are always one.

$$1 = E\left(mR^i\right) \tag{3.4}$$

In the special case of the riskfree asset:

$$1 = E\left(mR^f\right) \quad \Leftrightarrow \quad R^f = \frac{1}{E(m)} \tag{3.5}$$

Applying the covariance decomposition:

$$1 = E(m)E\left(R^i\right) + cov\left(m, R^i\right) \tag{3.6}$$

Using the correlation coefficient:

$$1 = E\left(mR^i\right) = E\left(m\right) E\left(R^i\right) - \rho_{m,R^i} \sigma\left(R^i\right) \sigma\left(m\right) \tag{3.7}$$

All assets priced by the stochastic discount factor m must obey:

$$\left|E\left(R^i\right) - R^f\right| \leq \frac{\sigma\left(m\right)}{E\left(m\right)} \sigma\left(R^i\right) \tag{3.8}$$

The Sharpe ratio is limited by the volatility of the discount factor.

$$\left|\frac{E\left(R\right) - R^f}{\sigma\left(R\right)}\right| \leq \frac{\sigma\left(m\right)}{E\left(m\right)} \tag{3.9}$$

If a return is mean-variance efficient, it lies on the mean-standard deviation frontier. The slope of the mean-standard deviatíon frontier is the highest available Sharpe ratio.

$$\left|\frac{E\left(R^{mv}\right) - R^f}{\sigma\left(R^{mv}\right)}\right| = \frac{\sigma\left(m\right)}{E\left(m\right)} \tag{3.10}$$

In the case of the consumption-based model with CRRA-utiliy,

$$\left|\frac{E\left(R^{mv}\right) - R^f}{\sigma\left(R^{mv}\right)}\right| = \frac{\sigma\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]}{E\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]} \tag{3.11}$$

if consumption growth is lognormal,

$$\left|\frac{E\left(R^{mv}\right) - R^f}{\sigma\left(R^{mv}\right)}\right| = \sqrt{e^{\gamma^2 \sigma^2 \left(\Delta \ln c_{t+1}\right)} - 1} \approx \gamma \sigma\left(\Delta \ln c_{t+1}\right). \tag{3.12}$$

Reading the equation[1], the slope of the mean-standard deviation frontier is higher if the economy is riskier - if consumption is more volatile - or if investors are more risk averse.

[1]See appendix A for the derivation of the first equality.

3.2 The Hansen-Jagannathan Bound

The equations (3.4) and (3.5) imply a discounted value of zero for an excess return defined as $R^e = R^i - R^f$,

$$0 = E\left(mR^e\right) = E\left(m\right) E\left(R^e\right) - \rho_{m,R^e}\sigma\left(R^e\right)\sigma\left(m\right). \tag{3.13}$$

Rearranging and using the fact that $|\rho| \leq 1$,

$$\frac{\sigma\left(m\right)}{E\left(m\right)} \geq \frac{\left|E\left(R^e\right)\right|}{\sigma\left(R^e\right)}. \tag{3.14}$$

Hansen and Jagannathan (1991) had the insight to read this equation as a restriction on the set of *discount factors* that can price a given set of returns, as well as a restriction on the set of *returns* given a specific discount factor. So we need very volatile discount factors with a mean near one to understand stock returns. This is the central tool in understanding and surmounting the equity premium puzzle which is discussed in the next chapter.

4 Performance of the Consumption-Based Model

4.1 Equity Premium Puzzle

The Hansen-Jagannathan bound implies

$$\frac{\sigma(m)}{E(m)} \geq \frac{|E(R^e)|}{\sigma(R^e)}. \tag{4.1}$$

One can show that in the consumption-based model

$$\frac{\sigma(m)}{E(m)} \approx \gamma\sigma(\Delta \ln c). \tag{4.2}$$

This relationship holds exactly in continuous time. So we can write

$$\gamma\sigma(\Delta \ln c) \geq \frac{|E(R^e)|}{\sigma(R^e)}. \tag{4.3}$$

The postwar mean value-weighed NYSE-return is about 8 percent per year over the T-bill rate, with a standard deviation of about 16 percent. Thus, the market Sharpe ratio $E(R^e)/\sigma(R^e)$ is about 0.5 for an annual investment horizon. A constant riskfree rate would determine

$$E(m) = \frac{1}{R^f}. \tag{4.4}$$

The T-bill rate is not very risky, so $E(m)$ is not far from the inverse of the mean T-bill rate, or about $E(m) \approx 0.99$. Thus, these basic facts about the mean and variance of stocks and bonds imply $\sigma(m) > 0.5$. The volatility of the stochastic discount factor must be about 50 percent of its level in annual data. Per capita consumption growth has a standard deviation about one percent per year. That implies $\sigma(m) = 0.01 = 1\%$. To match the equity premium we need $\gamma > 50$ and this is an implausibly high value of risk aversion.

4.2 Riskfree Rate Puzzle

Another way to state the equity premium puzzle is to plug in a serious value for γ and then to ask if the riskfree rate should be much higher. With lognormal consumption growth the riskfree rate[1] is

$$r_t^f = \delta + \gamma E_t \left(\Delta \ln c_{t+1}\right) - \frac{\gamma^2}{2}\sigma_t^2 \left(\Delta \ln c_{t+1}\right). \tag{4.5}$$

Using some estimated or assumed values taken over from Campbell and Cochrane (Campbell and Cochrane 1999), $\delta = 0.01$ and $\gamma = 50$ for the variables in that formula and :

$$r_t^f = 0.01 + 50 \cdot 0.0189 - \frac{50^2}{2} 0.015^2 = 0.67375. \tag{4.6}$$

The problem is that we do not observe a riskfree rate of 67.375% in the data.

4.3 Correlation Puzzle

The Hansen-Jagannathan bound takes the extreme possebility that consumption growth and stock return are perfectly correlated. In real data they are not perfectly correlated. The correlation of annual stock returns and nondurables and services consumption growth in postwar U.S. data is no more than about 0.2. If we use this information as well, the calculation becomes

$$\frac{\sigma\left(m\right)}{E\left(m\right)} \geq \frac{1}{\left|\rho_{m,R^e}\right|}\frac{\left|E\left(R^e\right)\right|}{\sigma\left(R^e\right)} = \frac{1}{0.2}0.5 = 2.5 \tag{4.7}$$

with $\sigma\left(m\right) \approx \gamma\sigma\left(\Delta c\right)$; we now need a risk aversion coefficient of 250. It is common to use risk aversion coefficients between one and five.

4.4 A Possible Solution

The conditional Hansen-Jagannathan bound implies:

$$\frac{E_t\left(R_{t+1}^e\right)}{\sigma_t\left(R_{t+1}^e\right)} = \rho_t\left(R_{t+1}^e, m_{t+1}\right)\frac{\sigma_t\left(m_{t+1}\right)}{E_t\left(m_{t+1}\right)} \tag{4.8}$$

[1]See Appendix B for the derivation of this formula.

In the case of the consumption-based model it would be[2]:

$$\sigma_t\left(m_{t+1}\right) = \gamma_t \sigma_t\left(\Delta c_{t+1}\right) \tag{4.9}$$

The Sharpe ratio varies over time. On the right-hand side, the conditional mean discount factor equals the inverse of the riskfree rate and so must be relatively stable over time. Time-varying conditional correlations are one possibility. The predictability of excess returns strongly suggests that the discount factor must be conditionally heteroskedastic. So $\sigma_t\left(m_{t+1}\right)$ must vary over time at least for the discount factor on the volatility bound, since $|\rho_{m,R^{mv}}| = 1$. For the consumption-based model as a special case consumption volatility does not vary over time so risk aversion must. Thus, we need either time-varying consumption risk or time-varying risk aversion and so a time-varying γ.

4.5 Long Horizon Predictability

For long horizons the Sharpe ratio can be written:

$$\frac{\left|E\left(R^e_{t\to t+k}\right)\right|}{\sigma\left(R^e_{t\to t+k}\right)} \leq \frac{\sigma\left(m_{t\to t+k}\right)}{E\left(m_{t\to t+k}\right)} \approx \gamma\sigma\left(\Delta c_{t\to t+k}\right) \tag{4.10}$$

If returns are predictable, this can imply mean-reversion and Sharpe ratios that rise faster than the square root of the horizon. Then the volatility of the discount factor must increase faster than the square root of the horizon. So the equity premium puzzle becomes even worse for long investment horizons.

[2] $\ln C = c$

5 Description of Data and Analysis

5.1 Consumption Data

Since the model uses additive separable utility, only consumption expenditures should be used which generate utility in the same period. This is true only for nondurable consumption goods and services.
Meyer (Meyer 1999 p.133) uses consumption expenditures on food, clothing and shoes, housing and energy, and other goods and services from January 1960 to December 1994 provided by the Deutsches Institut für Wirtschaftsforschung (DIW). There have also been data of nondurables and services by the Statistisches Bundesamt since 1980[1]. Meyer uses a weighed combination of the DIW-data to match this time series and extrapolates it for the time interval between 1960 and 1980. The time series C-Q contains quarterly data of total consumption growth and C-Y is the yearly equivalent. The time series CNS-Q contains quarterly data on consumption growth of expenditures on nondurables and services and again CNS-Y is the yearly equivalent.

C-Q mean %	0.695
C-Q Standard deviation %	0.846
CNS-Q mean %	0.562
CNS-Q standard deviation %	0.835
C-Y mean %	2.831
C-Y standard deviation %	2.077
CNS-Y mean %	2.303
CNS-Y standard deviation %	1.836

Table 5.1: Parameters according to Meyer

The following figures are my own estimates based on data of the Statistisches Bundesamt[2].

[1] Sum of the rows 33 (nondurables) and 38 (other services) of table 3.2.2.1, Fachserie 18, Reihe 1.3 "Privater Verbrauch nach Lieferbereichen und nach Dauerhaftigkeit und Wert der Güter"

[2] Quelle: Statis Datenbank: Zeitreihen Nr. 4303002, 4303003, 4303004, 4303006: Datei: `Consumption and Nondurables and states.xls`

PrVerb-Y mean %	2.427
PrVerb-Y standard deviation %	1.743

Table 5.2: Own Estimated Parameters

Because the standard deviation of consumption is lower as compared to the data of Meyer, it should be a harder target to explain the equity premium using these data.

5.2 Asset Price Data

Asset data are derived from the DAFOX[3]. This index is capital-weighed and constructed especially for research needs. It includes all German stocks traded in the first segment[4] at the Frankfurt Stock Exchange. The DAFOX is adjusted for dividends and the DAFOX-PR is not. The dividend on a share of the DAFOX can be calculated using the different performances of the two indices. The following table shows some statistics of the data for the period between 1960 and 1994[5]:

Stock Market Data

DAFOX annual return: mean %	7.113
DAFOX annual return: standard deviation %	14.099

Table 5.3: Statistics based on Log Growth Rates

As a proxy for the riskfree rate the Diskontsatz of the Deutsche Bundesbank is used[6]. To measure inflation and to calculate per capita values I have used data from the Statistisches Bundesamt[7]. Only West German data are included.

Inflation %	3.365
Risk free rate (Diskontsatz)%	4.71
3-month money market rate %	6.32[8]

Table 5.4: Inflation and riskfree rate

[3]Deutscher Aktien Forschungsindex

[4]Amtlicher Markt

[5]Datei: `stockdataauftagesbasis.xls`

[6]www.bundesbank.de

[7]Statis Datenbank: Zeitreihen Nr.: 1482001 (inflation) and 0035001, 0035002 (population data): Datei: `Consumption and Nondurables and states.xls`

5.3 Is there an Equity Premium Puzzle in Germany?

I use a sample from 1960 to 1994 and the standard consumption-based asset pricing model. Solving the following equation for γ and plugging in the estimated values delivers a solution.

$$\gamma\sigma(\Delta c) \geq \frac{|E(R^e)|}{\sigma(R^e)} \quad \Longleftrightarrow \quad \gamma \geq \frac{|E(R^e)|}{\sigma(R^e)} \frac{1}{\sigma(\Delta c)} \tag{5.1}$$

Using again $\rho = 1$ and data based on logarithmic differences:

$$\gamma = \frac{1}{0.01743} \frac{0.071126295 - 0.0471}{0.140989481} = 9.777 \tag{5.2}$$

Using a three month money market rate:

$$\gamma = \frac{1}{0.01743} \frac{0.071126295 - 0.0632}{0.140989481} = 3.225 \tag{5.3}$$

In all cases I have not considered a time varying riskfree interest rate because the habit formation model presented later on does not either. The period from 1960 to 1994 does not give a representative picture of a postwar data sample because of the stock market boom around 1960. The extreme high returns from 1949 to 1960 are not considered. Using the DAX-Rendite-Dreieck[9] the annual return was 33.5% for this period[10]. So, using a full postwar sample would increase γ even more. Using $\rho < 1$ to consider returns, which are not mean-variance efficient would also raise γ. So a γ of about 9.8 or 3.2 draws a flattering picture of the standard consumption-based asset pricing model. Is $\gamma = 3.225$ or $\gamma = 9.777$ too big? Using the formula for the riskfree rate with the German data (4.5), $\delta = 0.01$ and $\gamma = 3.225$ or $\gamma = 9.777$ one gets the follwing result:

$$r_t^f = 0.01 + 3.225 \cdot 0.02427 - \frac{3.225^2}{2} 0.0003038 = 0.08669 \tag{5.4}$$

$$r_t^f = 0.01 + 9.777 \cdot 0.02427 - \frac{9.777^2}{2} 0.0003038 = 0.23277 \tag{5.5}$$

Postwar data never show a real riskfree rate of 8.669% or 23.277%.

[9] Deutsches Aktieninstitut: www.dai.de

[10] Meyer 1999 p. 112 explains why the DAX can not be used

6 Habit Formation

6.1 The Habit Formation Model

There are some links between asset markets and macroeconomics. We expect the agents to smooth their consumption by holding assets. The introduction of financial markets allows risk trading. Agents can ensure states of low consumption growth or they can do risk taking and earn a risk premium by bying assets which pay of low in states of low consumption growth and pay of high in states of high consumption growth. The habit formation model links asset prices and consumption in a special way. Utility is a function of consumption and habit. The level of consumption is not important but how much consumption exceeds habit. Campbell and Cochrane use aggregate habit. So the individual history of consumption is unimportant.

6.1.1 Preferences and Technology

Identical agents maximize the utiliy function

$$E \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma}. \tag{6.1}$$

X_t is the level of habit and δ[1] is the subjective time-discount factor. The consumption surplus ratio is defined by

$$S_t \equiv \frac{C_t - X_t}{C_t}. \tag{6.2}$$

The local curvature η_t of the utility function is related[2] to S_t by

$$\eta_t \equiv -\frac{C_t u_{cc}(C_t, X_t)}{u_c(C_t, X_t)} = \frac{\gamma}{S_t}. \tag{6.3}$$

[1]formally known as β

[2]This is exactly the Arrow-Pratt coefficient of relative risk aversion. A derivation is given in appendix C.

Campbell and Cochrane use an *external habit* specification in which habit is determined by the history of *aggregate* consumption. The aggregate consumption surplus ratio is defined as

$$S_t^a \equiv \frac{C_t^a - X_t}{C_t^a} \tag{6.4}$$

where C_t^a denotes average consumption. How each individual's habit X_t responds to the history of aggregate consumption C_t^a is given by the following process for S_t^a. Using lower case letters to denote the logs of corresponding upper case letters, s_t^a evolves as a heteroskedastic first order autoregressive process[3],

$$s_{t+1}^a = (1-\phi)\,\bar{s} + \phi s_t^a + \lambda\left(s_t^a\right)\left(c_{t+1}^a - c_t^a - g\right). \tag{6.5}$$

ϕ, g and $\bar{s}$ are parameters and $\lambda\left(s_t^a\right)$ is the *sensitivity function.* In equilibrium, identical individuals choose the same level of consumption, so $C_t = C_t^a$, $S_t = S_t^a$. Near the steady state $s_t = \bar{s}$, the specification is approximately a standard linear specification in which habit responds slowly to consumption,

$$\begin{aligned} x_{t+1} &\approx \left[(1-\phi)\,h + g\right] + \phi x_t + (1-\phi)\,c_t \\ &\approx \left[h + \frac{g}{1-\phi}\right] + (1-\phi)\sum_{j=0}^{\infty}\phi^j c_{t-j} \end{aligned} \tag{6.6}$$

where $h = ln\left(1-\bar{S}\right)$ is the steady state of $x - c^4$. So the logarithm of habit is a slowly decaying moving average of the logarithm of consumption (Campbell and Cochrane 1995). Consumption growth follows an i.i.d. lognormal process,

$$\Delta c_{t+1} = g + v_{t+1}; \quad v_{t+1} \sim i.i.d. \quad N\left(0, \sigma^2\right). \tag{6.7}$$

If the last equation is regarded as an endowment process, the model can be closed as an endowment economy.

[3]I will sometimes use the common abbreviation AR(1) process.
[4]See appendix D for derivation.

6.1.2 Marginal Utility

Marginal utility is given by

$$u_c\left(C_t, X_t\right) = \left(C_t - X_t\right)^{-\gamma} = S_t^{-\gamma} C_t^{-\gamma}. \tag{6.8}$$

The intertemporal rate of substitution or stochastic discount factor is given by

$$M_{t+1} \equiv \delta \frac{u_c\left(C_{t+1}, X_{t+1}\right)}{u_c\left(C_t, X_t\right)} = \delta \left(\frac{S_{t+1}}{S_t} \frac{C_{t+1}}{C_t}\right)^{-\gamma}. \tag{6.9}$$

It is related to the state variable s_t and the log consumption innovation v_{t+1} by[5]

$$\begin{aligned} M_{t+1} &= \delta G^{-\gamma} e^{-\gamma\left(s_{t+1} - s_t + v_{t+1}\right)} \\ &= \delta G^{-\gamma} e^{-\gamma\left[(\phi-1)(s_t - \bar{s}) + (1+\lambda(s_t)) v_{t+1}\right]}. \end{aligned} \tag{6.10}$$

As derived in chapter three the conditional Sharpe ratio must obey

$$\frac{E_t\left(R_{t+1}^e\right)}{\sigma_t\left(R_{t+1}^e\right)} = -\rho_t\left(M_{t+1}, R_{t+1}^e\right) \frac{\sigma_t\left(M_{t+1}\right)}{E_t\left(M_{t+1}\right)} \leq \frac{\sigma_t\left(M_{t+1}\right)}{E_t\left(M_{t+1}\right)} \tag{6.11}$$

where ρ_t denotes conditional correlation. Since M is conditionally lognormal the largest possible Sharpe ratio can be found by[6]

$$\max_{\{all\ assets\}} \frac{E_t\left(R_{t+1}^e\right)}{\sigma_t\left(R_{t+1}^e\right)} = \left(e^{\gamma^2 \sigma^2 [1+\lambda(s_t)]^2} - 1\right)^{\frac{1}{2}} \approx \gamma \sigma \left[1 + \lambda\left(s_t\right)\right]. \tag{6.12}$$

The riskfree rate is given by[7]

$$r_t^f = -\ln\left(\delta\right) + \gamma g - \gamma\left(1 - \phi\right)\left(s_t - \bar{s}\right) - \frac{\gamma^2 \sigma^2}{2}\left[1 + \lambda\left(s_t\right)\right]^2. \tag{6.13}$$

The last term reflects precautionary savings. If uncertainty increases, people save more which drives down the riskfree rate.

[5]See appendix E for derivation.
[6]See appendix F for derivation.
[7]See appendix G for derivation.

6.1.3 Choosing the Sensitivity Function

The sensitivity function $\lambda(s_t)$ is chosen to satisfy three conditions. 1) The riskfree interest rate is constant. 2) Habit is predetermined at the steady state $s_t = \bar{s}$. 3) Habit is predetermined near the steady state, or, equivalently, habit moves positively with consumption everywhere. Habit cannot be completely predetermined in the model, or a suffiently low realization of consumption growth would leave consumption below habit, in which case a CRRA utility function is undefined. Hence, habit is predetermined at and near the steady state. These three considerations lead to a restriction that must hold between the consumption surplus ratio and the other parameters of the model:

$$\bar{S} = \sigma\sqrt{\frac{\gamma}{1-\phi}} \tag{6.14}$$

They lead to a specification of the sensitivity function of the following form:

$$\lambda(s_t) = \begin{cases} \left(1/\bar{S}\right)\sqrt{1-2(s_t-\bar{s})}-1 & , \quad s_t \leq s_{max} \\ 0 & , \quad s_t \geq s_{max} \end{cases} \tag{6.15}$$

The value at which the square root in the upper equation runs into zero is:

$$s_{max} \equiv \bar{s} + \frac{1}{2}\left(1-\bar{S}^2\right) \tag{6.16}$$

In the continuous time limit, the s_t process never reaches the region $s > s_{max}$. The riskfree rate is a constant[8],

$$r_t^f = -\ln\delta + \gamma g - \left(\frac{\gamma}{\bar{S}}\right)^2\frac{\sigma^2}{2} = -\ln\delta + \gamma g - \frac{\gamma}{2}(1-\phi). \tag{6.17}$$

Second, differentiating the transition equation (6.5), leads to[9]

$$\frac{dx_{t+1}}{dc_{t+1}} = 1 - \frac{\lambda(s_t)}{e^{-s_{t+1}}-1} \approx 1 - \frac{\lambda(s_t)}{e^{-s_t}-1}. \tag{6.18}$$

[8] See appendix H for derivation.
[9] See appendix I for derivation.

The latter approximation holds near the steady state. To obtain $dx/dc = 0$ at $s_t = \bar{s}$, a necessary requirement is:

$$\lambda(\bar{s}) = \left(1/\bar{S}\right) - 1 \tag{6.19}$$

Evaluating equation (6.15) at $\bar{s}$, it satisfies this condition. Third, to ensure that habit is predetermined in a neighborhood of the steady state

$$\left.\frac{d}{ds}\left(\frac{dx}{dc}\right)\right|_{s=\bar{s}} = 0 \tag{6.20}$$

must hold[10]. This condition also implies that habit moves positively with consumption everywhere, since dx/dc is a U-shaped function of s. Taking the derivative $d/ds\,(dx/dc)$ of the expression in (6.18) and setting it to zero at $s = \bar{s}$, Campbell and Cochrane obtain[11]

$$\lambda'(\bar{s}) = -\left(1/\bar{S}\right). \tag{6.21}$$

Equation (6.15) satisfies this condition. Since the functional form of $\lambda(s_t)$ has already been determined by the first two conditions, this condition determines the constraint (6.14) on the parameters of the model.

6.1.4 Prices of Long-Lived Assets

A stock is a claim to a consumption stream. From the basic pricing equation and the definition of returns,

$$1 = E_t\left[M_{t+1}R_{t+1}\right]; \quad R_{t+1} \equiv \frac{P_{t+1} + D_{t+1}}{P_t}, \tag{6.22}$$

the price-dividend or analogously the price-consumption ratio for a consumption claim satisfies[12]

$$\frac{P_t}{C_t}(s_t) = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t}\left(1 + \frac{P_{t+1}}{C_{t+1}}(s_{t+1})\right)\right]. \tag{6.23}$$

[10]See appendix J for derivation.
[11]A derivation is given in the NBER working-paper version Campbell and Cochrane (1995) p.12.
[12]See appendix K for a derivation.

The surplus consumption ratio s_t is the only state variable in this economy, so the price-consumption ratio is a function only of s_t. Campbell and Cochrane substitute M_{t+1} for (6.10) and consumption growth for (6.7) and solve this functional equation numerically on a grid for the state variable s_t, using numerical integration over the normally distributed shock v_{t+1} to evaluate the conditional expectation.

$$\frac{P}{C}\left(s_t\right) = \delta G^{1-\gamma} E\left[e^{-\gamma\left(s_{t+1}-s_t\right)} e^{(1-\gamma)v_{t+1}} \left(1 + \frac{P}{C}\left(s_{t+1}\right)\right) \middle| s_t \right] \tag{6.24}$$

Plugging in the expression for $s_{t+1} - s_t$:

$$\frac{P}{C}\left(s_t\right) = \delta G^{1-\gamma} E\left[e^{-\gamma\left[(1-\phi)(\bar{s}-s_t)+\lambda(s_t)v_{t+1}\right]+(1-\gamma)v_{t+1}} \left(1 + \frac{P}{C}\left(s_{t+1}\right)\right) \middle| s_t \right] \tag{6.25}$$

This can be rearranged to:

$$\frac{P}{C}\left(s_t\right) = \delta G^{1-\gamma} e^{-\gamma\left[(1-\phi)(\bar{s}-s_t)\right]} E\left[e^{\{-[(\lambda(s_t)+1)\gamma-1]\}v_{t+1}} \left(1 + \frac{P}{C}\left(s_{t+1}\right)\right) \middle| s_t \right] \tag{6.26}$$

Since the growth rates of stock market dividends and consumption are only weakly correlated in US data it makes sense to model dividends and consumption separately. The process for dividend growth takes the following form:

$$\Delta d_{t+1} = g + w_{t+1};\ w_{t+1} \sim i.i.d. \quad N\left(0, \sigma_w^2\right),\ corr\left(w_t, v_t\right) = \rho \tag{6.27}$$

The price-dividend ratio of a claim to the dividend stream then satisfies:

$$\frac{P_t}{D_t}\left(s_t\right) = E_t\left[M_{t+1} \frac{D_{t+1}}{D_t} \left(1 + \frac{P_{t+1}}{D_{t+1}}\left(s_{t+1}\right)\right)\right] \tag{6.28}$$

Plug in the dividend growth process:

$$\frac{P_t}{D_t}\left(s_t\right) = G E_t\left[M_{t+1} e^{w_{t+1}} \left(1 + \frac{P_{t+1}}{D_{t+1}}\left(s_{t+1}\right)\right)\right] \tag{6.29}$$

Using the Law of Iterated Expectations to avoid integrating over two random variables, v and w:

$$E\left[g\left(v\right)e^w\right] = E\left[E\left(g\left(v\right)e^w \mid v\right)\right] = E\left[g\left(v\right)E\left(e^w \mid v\right)\right] \tag{6.30}$$

Then the dividend claim satisfies:

$$\frac{P_t}{D_t}(s_t) = Ge^{(1/2)\left(1-\rho^2\right)\sigma_w^2} E_t\left[M_{t+1}e^{\rho(\sigma_w/\sigma)v_{t+1}}\left(1+\frac{P_{t+1}}{D_{t+1}}(s_{t+1})\right)\right] \tag{6.31}$$

Now substitute M_{t+1} in terms of the shock v_{t+1}:

$$\begin{aligned}\frac{P_t}{D_t}(s_t) &= Ge^{(1/2)\left(1-\rho^2\right)\sigma_w^2} \\ &\quad E_t\left[e^{\ln\delta-\gamma g-\gamma[(\phi-1)(s_t-\bar{s})+(1+\lambda(s_t))v_{t+1}]}e^{\rho(\sigma_w/\sigma)v_{t+1}}\left(1+\frac{P_{t+1}}{D_{t+1}}(s_{t+1})\right)\right]\end{aligned} \tag{6.32}$$

Now I plug in the expression for $\lambda(s_t)$:

$$\begin{aligned}\frac{P_t}{D_t}(s_t) &= Ge^{(1/2)\left(1-\rho^2\right)\sigma_w^2+\ln\delta-\gamma[g+(\phi-1)(s_t-\bar{s})+1]} \\ &\quad E_t\left[e^{\left\{-\gamma\left[\left(1/\bar{S}\right)\sqrt{1-2(s_t-\bar{s})}-1\right]+\rho(\sigma_w/\sigma)\right\}v_{t+1}}\left(1+\frac{P_{t+1}}{D_{t+1}}(s_{t+1})\right)\right]\end{aligned} \tag{6.33}$$

Now I use the following equation:

$$E_t\left(e^z\right) = e^{E_t(z)+(1/2)\sigma_t^2(z)} \tag{6.34}$$

$$\begin{aligned}\frac{P_t}{D_t}(s_t) &= Ge^{(1/2)\left(1-\rho^2\right)\sigma_w^2+\ln\delta-\gamma[g+(\phi-1)(s_t-\bar{s})+1]} \\ &\quad e^{(1/2)\left\{-\gamma\left[\left(1/\bar{S}\right)\sqrt{1-2(s_t-\bar{s})}-1\right]+\rho(\sigma_w/\sigma)\right\}^2\sigma^2} \\ &\quad E_t\left[e^{\left\{-\gamma\left[\left(1/\bar{S}\right)\sqrt{1-2(s_t-\bar{s})}-1\right]+\rho(\sigma_w/\sigma)\right\}v_{t+1}}\frac{P_{t+1}}{D_{t+1}}(s_{t+1})\right]\end{aligned} \tag{6.35}$$

This equation is solved numerically on a grid for the state variable s_t, using a numerical integrator to evaluate the conditional expectation over the normally distributed v_{t+1}.

6.1.5 An Analytical Solution for the Nonstochastic Case

The price-consumption ratio of an asset that lives infinitely long can be written as follows[13]:

$$\frac{P_t}{C_t} = E_t \sum_{j=1}^{\infty} \delta^j \frac{u'(C_{t+j})}{u'(C_t)} \frac{C_{t+j}}{C_t} = E_t \sum_{j=1}^{\infty} \delta^j \left(\frac{S_{t+j}}{S_t}\right)^{-\gamma} \left(\frac{C_{t+j}}{C_t}\right)^{1-\gamma} \tag{6.36}$$

For the nonstochastic case with $\sigma^2 = 0$ and $c_{t+1} = g + c_t$ this equation changes to:

$$\frac{P_t}{C_t} = \sum_{j=1}^{\infty} \left(\delta G^{1-\gamma}\right)^j e^{-\gamma(s_{t+j}-s_t)} = \sum_{j=1}^{\infty} \left(\delta G^{1-\gamma}\right)^j e^{\gamma\left(1-\phi^j\right)(s_t-\bar{s})} \tag{6.37}$$

6.1.6 Choosing Parameters

Campbell and Cochrane compare the model to two data sets. The first one contains postwar data from 1947 to 1995. They use the CRSP value-weighed NYSE stock index return, three-month treasury bill rate, and per-capita nondurables and services consumption. The second one is a century-long annual data set of S&P 500 stock and commercial paper returns (1871-1993) and per-capita consumption (1889-1992) from Campbell (Campbell 1998). Table 6.1 presents statistics of the data.

Assumed and Implied Parameters

Mean consumption growth* g,%	1.89
Standard deviation of consumption growth* σ,%	1.50
Log riskfree rate* r^f,%	0.94
Persistence coefficient* ϕ	0.87
Utility curvature γ	2.00
Standard deviation of dividend growth* σ_w,%	11.2
Correlation between Δd and Δc, ρ	0.02

Subjective discount factor* δ	0.89
Steady state surplus consumption ratio $\bar{S}$	0.057
Maximum surplus consumption ratio S_{max}	0.094

Starred (*) values are annualized values, e.g. 12g, $\sqrt{12}\sigma$, $12r^f$, ϕ^{12}, δ^{12} since the model is simulated at a monthly frequency.

Table 6.1: Parameter choices

[13] See appendix L for derivation.

Some free parameters of the model are chosen to match the postwar data. Mean and standard deviation of log consumption growth, g and σ, are chosen to match the consumption data. The serial correlation parameter ϕ is chosen to match the serial correlation of log price-dividend ratios. The subjective discount factor δ matches the riskfree rate with the average real return on Treasury bills. The parameter of risk aversion γ matches ratio of unconditional mean to unconditional standard deviation of excess returns from artificial data to real data. The standard deviation σ_w is taken from CRSP data. The choice for ρ is tricky. If dividend growth was uncorrelated with consumption growth, a claim for dividend growth would have no risk premium. Correlations are difficult to measure. Campbell and Cochrane choose a correlation of 0.2. The results are robust in relation to the precise value of this correlation.

6.1.7 Solution and Evaluation

First the model is solved numerically. Second Campbell and Cochrane simulate data by drawing shocks from a random number generator and compare the artificial data to the real data. Finally they feed the model historical consumption shocks to see what it tells about historical movements in stock prices.

Asset Prices and the Surplus Consumption Ratio

Stationary distribution of the surplus consumption ratio

Figure 6.1 shows the density of the surplus consumption ratio. The figure shows the distribution of the continuous time version for the process, calculated in the appendix of Campbell and Cochrane (Campbell and Cochrane 1998b).

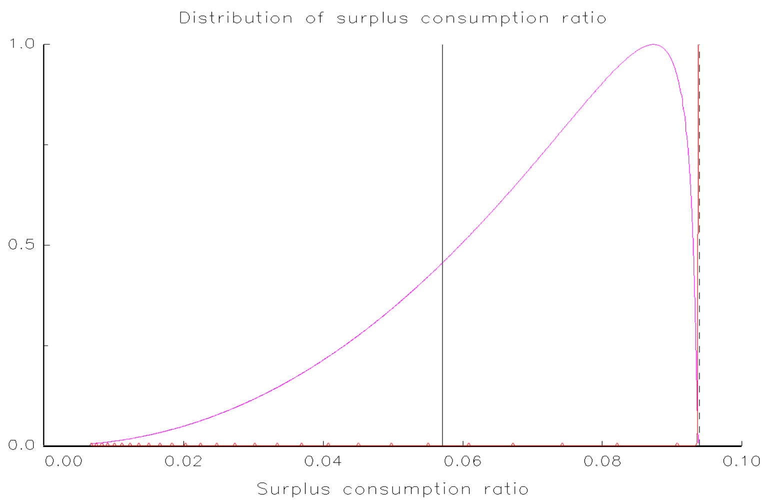

Figure 6.1: Density of s

Figure 6.2 shows the sensitivity function λ which is used to take hetereskedasticity into the process of s. Figure 6.3 shows how habit responds to changes of consumption. At the steady state there is no response. Campbell and Cochrane refer to surplus consumption ratio below the steady state level $\bar{s}$ as a "recession" and beyond as a "boom".

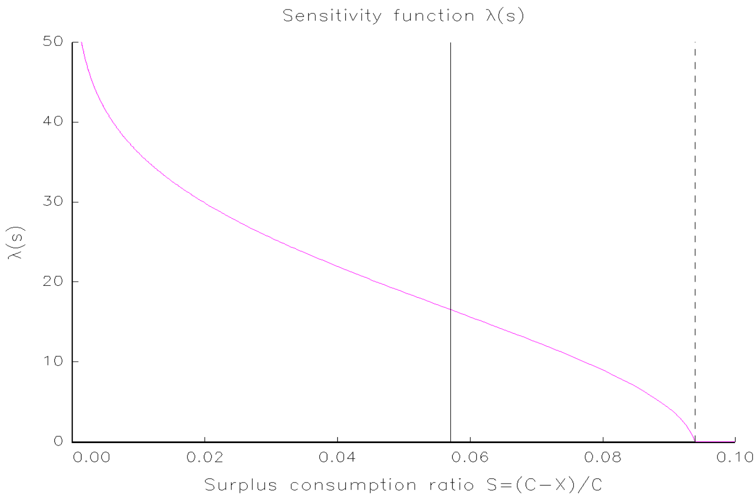

Figure 6.2: Sensitivity function λ(s)

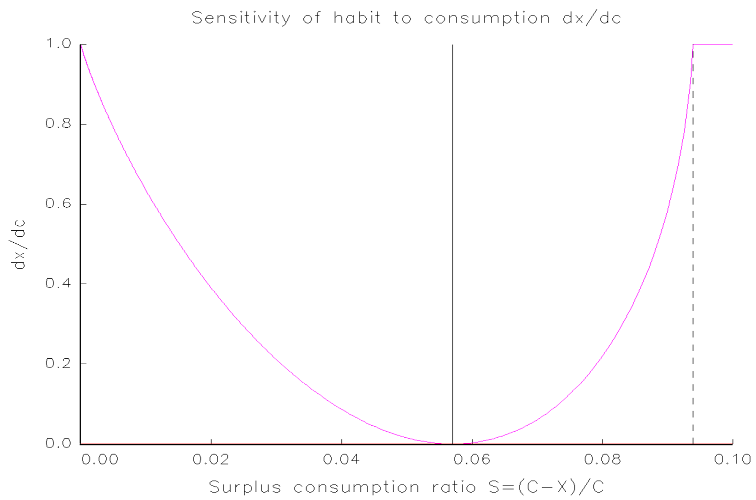

Figure 6.3: Sensitivity of habit to consumption

Price-dividend ratios and the surplus consumption ratio

Figure 6.4 plots the price-dividend ratio of the consumption claim and the dividend claim against the surplus consumption ratio. Both increase with the surplus consumption ratio. When consumption is low relative to habit in a recession, the curvature of the utility function is high, and prices are depressed relative to dividend and consumption. The price-dividend ratio and price-consumption ratio are almost

exactly the same despite the low correlation of 0.2. Dividend growth is much more volatile than consumption growth, so the regression coefficient $\beta = \rho\sigma_{\Delta d}/\sigma_{\Delta c}$ of dividend growth on consumption growth is roughly one. The priced components of the two assets are similar, and therefore are their prices.

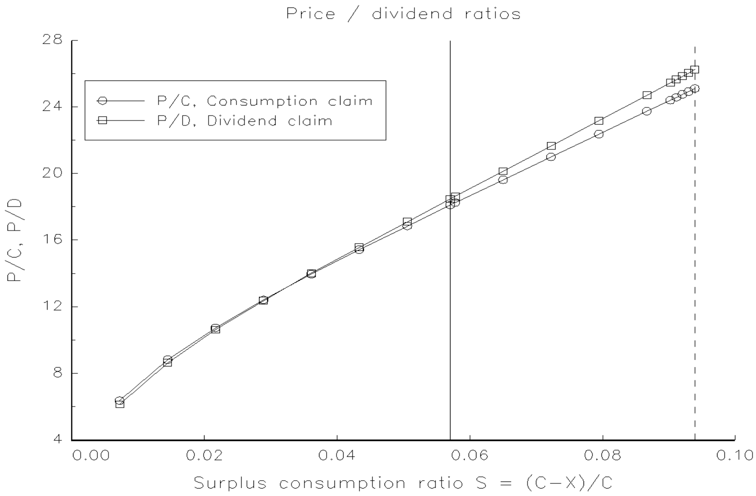

Figure 6.4: Price-dividend ratios

Conditional moments of returns

Figure 6.5 plots the expected consumption-claim and dividend-claim returns and the riskfree interest rate against the surplus consumption ratio. As consumption declines towards habit, expected returns rise dramatically over the constant riskfree rate.

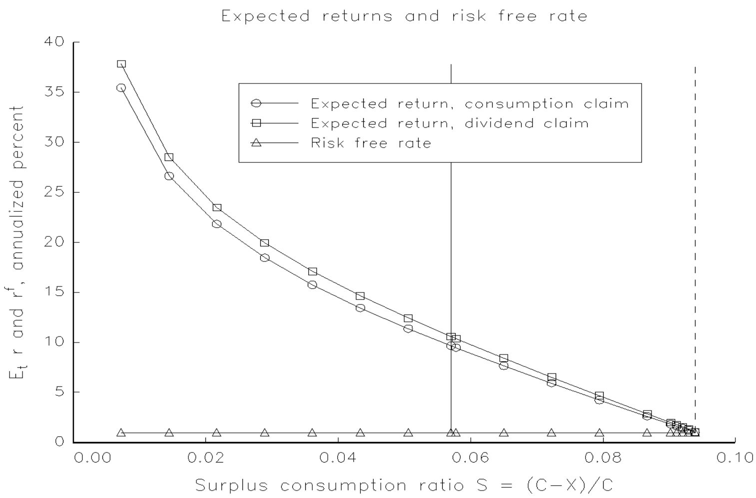

Figure 6.5: Expected returns

Figure 6.6 presents the conditional standard deviations of returns as functions of the surplus consumption ratio. As consumption declines towards habit, the conditional

variance of returns increases. The model produces several effects that have been emphasized in the ARCH literature: highly autocorrelated conditional variance in stock returns by Tim Bollerslev (Bollerslev et. al 1992), a leverage effect by Fisher Black (Black 1976) which means that price declines increase volatility, and countercyclical variation in volatility. The return on the dividend claim is

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{P_{t+1}/D_{t+1} + 1}{P_t/D_t} \times \frac{D_{t+1}}{D_t}. \tag{6.38}$$

The expected returns are nearly identical because the price-dividend ratio and price-consumption ratio are nearly identical functions of state, and dividend and consumption growth are not predictable. The conditional standard deviation of the dividend claim inherits the same dependence on state through the nearly identical P/D term, but adds the extra, constant, standard deviation of dividend growth.

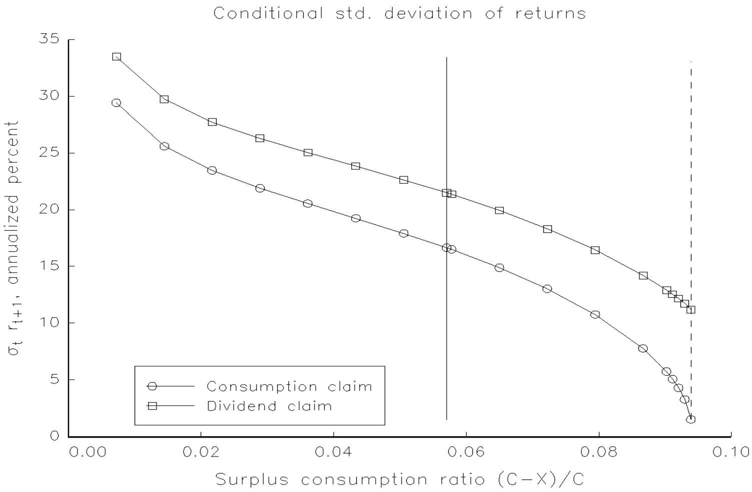

Figure 6.6: Conditional standard deviations of returns

Conditional Sharpe ratios

Figure 6.7 plots the Sharpe ratio against the surplus consumption ratio. The top line is the maximum possible Sharpe ratio calculated from the Hansen-Jagannathan bound. The consumption claim nearly attains the Sharpe ratio bound, implying that it is nearly conditionally mean-variance efficient. The consumption-claim model has only one shock. Hence the only reason the consumption claim (or any claim whose return depends on the single shock) is not exactly conditionally mean-variance efficient is that it is nonlinearly related to the shock. For the consumption claim, the effects of such nonlinearity are slight. The dividend claim has a slightly higher mean return and a substantially higher standard deviation, since there is a second dividend growth shock as well as the consumption (discount-rate) shock. The top line of Figure 6.7 is also interesting as a characterization of the discount factor. The

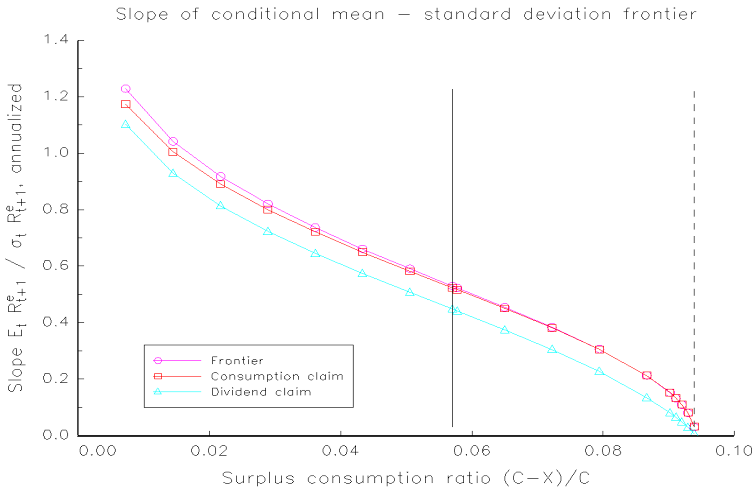

Figure 6.7: Sharpe ratios as functions of the surplus consumption ratio

conditional mean of the discount factor is constant, so this line plots the conditional standard deviation of the discount factor.

Statistics from Simulated Data

Campbell and Cochrane simulate 500,000 months of artificial data to calculate population values for a variety of statistics.

Statistic	Consumption claim	Dividend claim	Postwar data	Long data
$E(\Delta c)$	1.89*		1.89	1.72
$\sigma(\Delta c)$	1.22*		1.22	3.32
$E(r^f)$	0.94*		0.94	2.92
$E(r-r^f)/\sigma(r-r^f)$	0.43*	0.33	0.43	0.22
$E(R-R^f)/\sigma(R-R^f)$	0.50		0.50	
$E(r-r^f)$	6.64	6.52	6.69	3.90
$\sigma(r-r^f)$	15.2	20.0	15.7	18.0
$exp[E(p-d)]$	18.3	18.7	24.7	21.1
$\sigma(p-d)$	0.27	0.29	0.26	0.27

Table 6.2: Means and standard deviations of data

The model is simulated at a monthly frequency. Statistics are calculated from artificial time-averaged data at an annual frequency. Asterisks denote statistics that model parameters were chosen to replicate. All returns are annual percentages.

In order to facilitate a comparison with historical data, they simulate the model at a monthly frequency, and then construct time-averaged artificial annual data. As in the actual data, they average the level in each year. They form annual returns by

taking the product of intervening monthly returns.

Means and standard deviations

Table 6.2 presents means and standard deviations in simulated data, with the corresponding statistics from the two historical data sets. The first four moments are chosen to match the corresponding postwar statistics exactly. To match the Sharpe ratio for log returns in postwar data γ equals two. The model can match the mean and standard deviation of excess stock returns, with a constant low interest rate and a discount factor $\delta = 0.89$ less than one, by *any* choice of parameters.

Autocorrelations and cross-correlations

Table 6.3 presents autocorrelations in the data. The parameter ϕ is picked to generate the 0.87 first order annual autocorrelation of the price-dividend ratio seen in the table.

		Lag				
Variable	Source	1	2	3	5	7
$p-d$	Cons. claim	0.87	0.76	0.66	0.51	0.39
	Dividend claim	0.87	0.76	0.66	0.51	0.39
	Postwar data	0.87	0.77	0.70	0.41	0.04
	Long data	0.78	0.57	0.50	0.32	0.29
$r-r^f$	Cons. claim	-0.06	-0.05	-0.04	-0.02	-0.02
	Dividend claim	-0.05	-0.04	-0.03	-0.02	-0.01
	Postwar data	-0.11	-0.28	0.15	0.02	0.10
	Long data	0.05	-0.21	0.08	-0.14	0.11
$\sum_{i=1}^{j}\rho\left(r_t^e, r_{t-j}^e\right)$	Cons. claim	-0.06	-0.11	-0.15	-0.20	-0.26
	Dividend claim	-0.05	-0.09	-0.12	-0.14	-0.18
	Postwar data	-0.11	-0.39	-0.24	0.18	0.13
	Long data	0.05	-0.16	-0.09	-0.28	-0.15
$\lvert r\rvert$	Cons. claim	0.09	0.09	0.09	0.07	0.05
	Dividend claim	0.05	0.05	0.05	0.04	0.03
	Postwar data	0.08	-0.26	-0.10	-0.08	0.05
	Long data	0.13	0.09	0.07	0.14	0.15

Table 6.3: Autocorrelations of data

The model values are based on time-aggregated annual values using a monthly simulation interval. Data are all annual. The third block of numbers is the partial sum of return autocorrelations up to lag j.

Returns display a series of small negative autocorrelations that generate univariate mean reversion (Fama and French 1988a). The negative autocorrelations of returns also generate observations that price changes tend to be reversed. Since individual long-term autocorrelations are small and poorly measured, the partial sum of autocorrelation coefficients is shown in the table. The dividend claim has a lower autocorrelation of absolute returns. The cross-correlation between the price-dividend ratio and subsequent excess returns in table 6.4 shows that price-consumption ratio forecasts longhorizon returns, with the right sign: high prices forecast low returns and the forecastability of returns increases with the horizon. The cross-correlation between the price-dividend ratio or returns and subsequent absolute returns show that a low price-consumption ratio or a big price decline signal high volatility for many quarters ahead. This is the leverage effect mentioned by Black (Black 1976) and many others.

Long-horizon regressions

Table 6.5 presents long-horizon regressions of log excess stock returns on the log price-dividend ratio in simulated and historical data.

		Lag				
Variable	Source	1	2	3	5	7
$p_t - d_t, r^e_{t+j}$	Cons. claim	-0.35	-0.30	-0.26	-0.20	-0.15
	Dividend claim	-0.28	-0.24	-0.20	-0.16	-0.12
	Postwar data	-0.42	-0.25	-0.13	-0.35	-0.17
	Long data	-0.20	-0.21	-0.10	-0.19	-0.08
$r^e_t, \lvert r^e_{t+j} \rvert$	Cons. claim	-0.09	-0.07	-0.06	-0.03	-0.03
	Dividend claim	-0.06	-0.04	-0.04	-0.03	-0.02
	Postwar data	-0.32	-0.14	0.10	-0.04	-0.08
	Long data	-0.15	0.03	0.12	0.02	-0.01
$p_t - d_t, \lvert r^e_{t+j} \rvert$	Cons. claim	-0.49	-0.42	-0.37	-0.28	-0.21
	Dividend claim	-0.36	-0.31	-0.27	-0.21	-0.16
	Postwar data	-0.16	0.09	0.11	-0.05	0.02
	Long data	-0.12	0.02	-0.06	-0.10	-0.05

Table 6.4: Cross-correlations of data

It shows the classical pattern documented by Fama and French (Fama and French 1988b) and others. The coefficients are negative; so high prices imply low expected returns. The coefficients increase linearly with the horizon at first, but then less quickly. The R^2 starts low, but then rises. Note the lower R^2 for the dividend claim. Here the extra noise comes in. Fama (Fama 1990) interprets similar correlations of

returns with output as evidence that returns move on news of future cash flows. The stock markets anticipate high consumption growth.

Horizon (Years)	C. claim 10×coef.	D. claim 10×coef.	Postwar data 10×coef.	Long data 10×coef.
1	-2.0	-1.9	-2.6	-1.3
2	-3.7	-3.6	-4.3	-2.8
3	-5.1	-5.0	-5.4	-3.5
5	-7.5	-7.3	-9.0	-6.0
7	-9.4	-9.2	-12.1	-7.5
	C. claim R^2	D. claim R^2	Postwar data R^2	Long data R^2
1	0.13	0.08	0.18	0.04
2	0.23	0.14	0.27	0.08
3	0.32	0.19	0.37	0.09
5	0.46	0.26	0.55	0.18
7	0.55	0.30	0.65	0.23

Table 6.5: Long-horizon return regressions

The correlation of consumption growth with stock returns

In the model of Campbell and Cochrane, the unconditional correlation between monthly consumption claim returns and monthly consumption growth is 0.93. This value corresponds to the left hand panel of Figure 6.8.

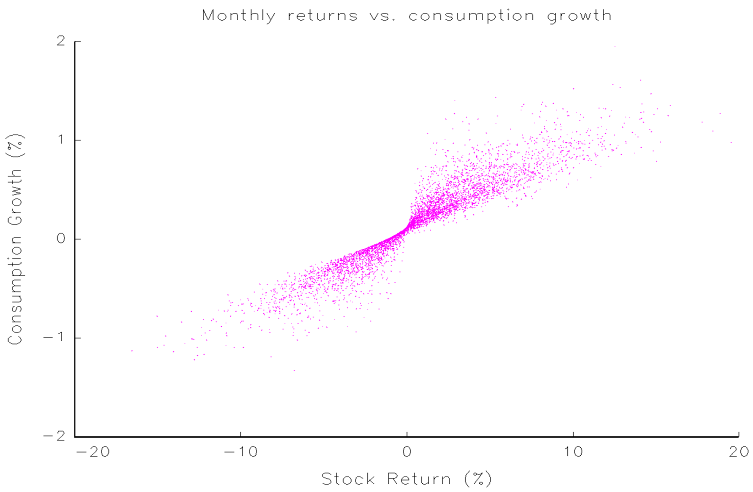

Figure 6.8: Simulated monthly consumption growth vs. monthly consumption claim returns

There is no correlation at any lead or lag. When the artificial data are aggregated to annual frequencies, the contemporaneous correlation drops to 0.47. Furthermore, time-aggregation produces a strong positive correlation between returns and subsequent consumption growth, and a negative correlation between returns and previous consumption growth. This is the same sign pattern that is in the data.

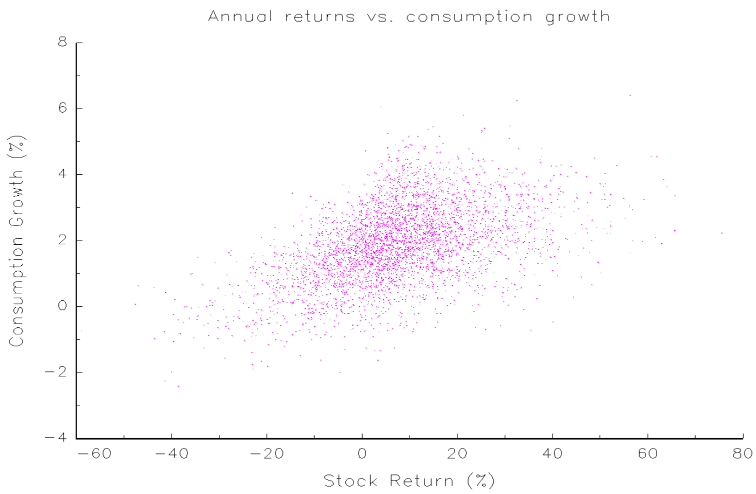

Figure 6.9: Growth in simulated annual consumption vs. annual consumption claim returns

	monthly		annual		Data		
Corr.	Cons. claim	Div. claim	Cons. claim	Div. claim	PW quarterly	PW ann.	Long ann.
$r^e_t, \Delta c_{t-2}$	0.0	0.0	-0.16	-0.13	-0.05	-0.16	-0.05
$r^e_t, \Delta c_{t-1}$	0.0	0.0	-0.19	-0.15	-0.10	-0.34	-0.08
$r^e_t, \Delta c_t$	0.93	0.79	0.47	0.40	0.12	-0.05	0.09
$r^e_t, \Delta c_{t+1}$	0.0	0.0	0.50	0.42	0.19	0.37	0.49
$r^e_t, \Delta c_{t+1}$	0.0	0.0	0.0	0.0	0.15	-0.26	0.05

Table 6.6: Correlation between consumption growth and stock return

The correlation of the discount factor with consumption growth and stock returns

Table 6.7 shows correlations. In monthly artificial data, the consumption claim return is far better correlated with the true discount factor than is consumption growth. Although the discount factor is *conditionally* perfectly correlated with consumption growth, the *unconditional* correlation is low, because the surplus consumption ratio varies. The relative performance of the consumption-based model should drop further at long horizons. At longer horizons, there is more movement of the surplus consumption ratio independend of consumption growth, and this movement will be revealed by stock return variation since stock prices decline when the surplus con-

sumption ratio declines. Time aggregation further obscures the signal. At monthly horizon, the dividend claim is a poorer proxy than even consumption growth, because dividend growth contains noise not correlated with the discount factor.

	Correlation of the discount factor with		
	Consumption growth	Consumption claim return	Dividend claim return
Monthly	0.90	0.99	0.83
Annual	0.45	0.99	0.80

Table 6.7: Correlation between the discount factor $M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \frac{S_{t+1}}{S_t} \right)^{-\gamma}$ and consumption growth, consumption claim return and dividend claim return.

Interpreting Historical Consumption and Stock Price Data

Campbell and Cochrane now feed the model actual instead of artificial data. Figure 6.10 shows how habit responds smoothly to changes in consumption, trending up in the high growth 60's and growing more slowly in the 70's. Cyclical dips in consumption bring consumption closer to habit. The model predicts low price-dividend ratios and high expected returns for those periods.

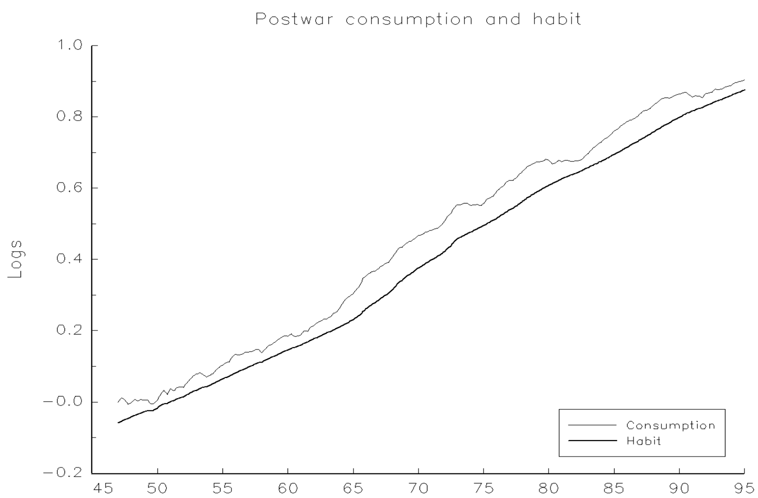

Figure 6.10: Nondurables and services consumption and habit level implied, assuming that s starts at $\bar{s}$

Figure 6.11 presents the model's predictions for the price-dividend ratio of a consumption claim, together with the actual price-dividend ratio on the S&P-500 index. The prewar prediction is based on a calibration of the model to the long data set; it uses the lower mean and higher standard deviation of consumption growth of these data and a lower value of $\gamma = 0.7$ to generate the lower Sharpe ratio in that data

set. When consumption declines for several years in a row, coming nearer to the constructed habit, stock prices fall. The decline in consumption in the Great Depression was so extreme that the model predicts an even larger fall in stock prices than actually occured. The worst performance occurs in the last few years. Possible excuses a shift in corporate financial policy towards the repurchase of equity rather than dividend payments; an increase in the consumption of stock-market investors that is not properly captured in the aggregate consumption data, due perhaps to rising income inequality in the period or the demographic effects of the baby boom generation entering peak saving years; and measurement problems such as compositional shifts of consumption away from nondurables and services[14].
Table 6.8 shows statistics for the data used in figure 6.11[15]. The average of historical price-dividend ratios would have a better match with the model. Price-dividend ratios are positive but weakly correlated. The log growth rates are nearly uncorrelated.

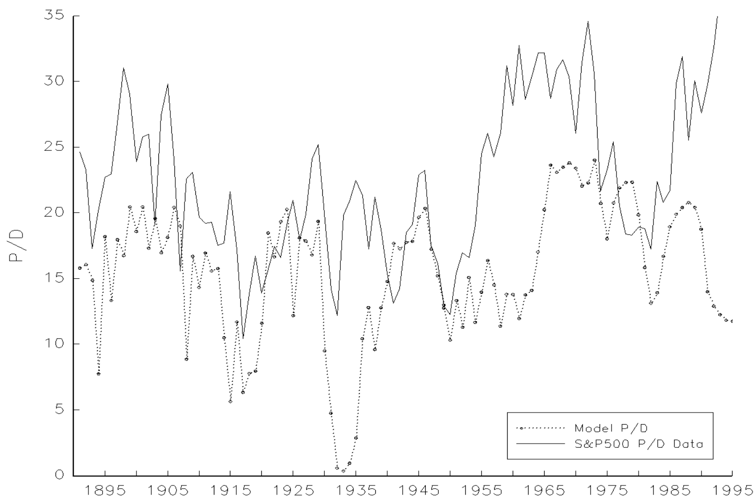

Figure 6.11: Historical price-dividend ratio, and model prediction based on the history of consumption

Model intuition

The equity premium and riskfree rate puzzles

The model is consistent with the equity premium and a low and constant riskfree rate. With power utility $M_{t+1} = \delta \left(C_{t+1}/C_t \right)^{-\gamma}$, a constant riskfree rate, and i.i.d.

[14]This argument contradicts the fact that the sector of services tends to grow fastest in a post-industrial economy.
[15]This is not part of the original paper.

Mean values

	hist. data	model data
$\frac{1}{T}\sum_{t=1}^{T}\frac{P_t}{D_t}$	15.76	22.84

Measures to evaluate the performance

$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{P_t^M}{D_t^M}-\frac{P_t^H}{D_t^H}\right)^2}$	9.83
$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{P_t^H}{D_t^H}-\frac{1}{T}\sum_{t=1}^{T}\frac{P_t}{D_t}\right)^2}$	6.28
$\rho\left(\frac{P_t^M}{D_t^M},\frac{P_t^H}{D_t^H}\right)$ $\rho\left(\Delta\left(p-d\right)_{t+1}^M,\Delta\left(p-d\right)_{t+1}^H\right)$	0.3 -0.01

Table 6.8: Statistics of data used in Figure 6.11

lognormal consumption growth with mean g and standard deviation σ, the Hansen-Jagannathan or Sharpe ratio inequality specializes to

$$\frac{E\left(R^e\right)}{\sigma\left(R^e\right)} \leq \sqrt{e^{\gamma^2\sigma^2}-1} \approx \gamma\sigma \tag{6.39}$$

and the log riskfree rate is

$$r_t^f = -\ln\left(\delta\right) + \gamma g - \gamma^2\frac{\sigma^2}{2} \tag{6.40}$$

From equation 6.39 follows that one needs a risk aversion coefficient $\gamma \geq 41$ to explain a Sharpe ratio of 0.50 with $\sigma = 1.22\%$. More importantly, a high value of γ makes the term γg in the riskfree rate equation very large. Thus $\gamma = 41$, $g = 1.89$ means one needs $\delta = 1.90$ to get a one percent riskfree rate. Imposing $\delta \leq 1$, one predicts

a riskfree rate of more than 90 percent per year. Setting $\delta = 1.90$ and $\gamma = 41$ the equation for the riskfree rate implies that the riskfree rate should be quite sensitive to the mean consumption growth rate, which is not the case. The model also features high curvature of the utility function. Though the power $\gamma = 2$ is low, the surplus consumption ratio is also low, so local curvature $\gamma = -Cu_{cc}/u_c = \gamma/S$ is high, 35 at the steady state, and higher in low surplus consumption ratio states.
Equation (6.16)

$$r_t^f = -\ln\delta + \gamma g - \left(\frac{\gamma}{\bar{S}}\right)^2 \frac{\sigma^2}{2}$$

shows that the parameter of risk aversion $\gamma = 2$, much lower than utility curvature γ/S, controls the relationship between average consumption growth and the riskfree interest rate.
This is the way a riskfree rate puzzle has been avoided: an intuively plausible $\delta = 0.89 < 1$ is consistent with the low observed real interest rate; and the model predicts a much less sensitive relationship across countries or over time between mean consumption growth and interest rate.
Abel (Abel 1998) highlights the danger for accounting for an equity premium by a term premium. Since interest rates are constant in the model, long-term bonds earn exactly the same returns as short-term bonds, and the entire equity premium is a risk premium, not a term premium.

The long-run equity premium

Consumption is roughly a random walk at any horizon, so the standard deviation of consumption growth grows roughly with the square root of the horizon[16]. The negative autocorrelation of stock returns means that k-year return variances are somewhat less than k times one year return variances, so the market Sharpe ratio grows if anything *faster* than the square root of the horizon[17]. In the model the k-period stochastic discount factor is

$$M_{t,t+k} = \delta^k \left(\frac{S_{t+k}}{S_t}\frac{C_{t+k}}{C_t}\right)^{-\gamma}. \tag{6.41}$$

Equation (6.11) implies that the standard deviation of this discount factor must increase roughly with the square root of the horizon to be consistent with the long-run equity premium, and even faster to generate the negative autocorrelation of stock returns. The model has a pure random walk in consumption, yet it produces negative autocorrelation in returns and therefore high Sharpe ratios at all horizons although

[16]Appendix M
[17]Appendix N

the state variable S_t is stationary. The clue is: while S_t is stationary, $S_t^{-\gamma}$ is not. S_t has a fat tail approaching zero, so the conditional variance of $S_{t+k}^{-\gamma}$ grows without bound. As $s \to -\infty$, the leading terms in the distribution are

$$f\left(s\right) \approx e^{-\gamma|s|-2\gamma\bar{S}\sqrt{2|s|}}. \tag{6.42}$$

s is a covariance-stationary process with a well-defined unconditional mean, variance and all higher moments. $S = e^s$ is also well-behaved. However, while $S^{-\gamma}$ has a finite unconditional mean, since $e^{-\gamma s} f\left(s\right)$ is integrable, $S^{-\gamma}$ does not have a finite unconditional variance since $e^{-2\gamma s} f(s) \approx e^{\gamma|s|}$ explodes as $s \to \infty$. The distinction between stationary S and nonstationary $S^{-\gamma}$ is central. Any model that wishes to explain the equity premium at long and short horizons runs by means of an additional, stationary state variable must find some similar transformation so that the equity premium remains high at long horizons.

A recession state variable

This model makes a fundamental change in the way we have to understand risk premia as emphasized by Equation (6.41). Consumers do not fear stocky because of the resulting risk to wealth per se; they fear stocks because stocks are likely to do poor in recessions, times of low consumption surplus ratios. While $(C_{t+1}/C_t)^{-\gamma}$ and $(S_{t+1}/S_t)^{-\gamma}$ enter symmetrically in the formula, the volatility of $(C_{t+1}/C_t)^{-\gamma}$ is so low that it accounts for essentially no risk premia. So variation in $(S_{t+1}/S_t)^{-\gamma}$ is much larger, and accounts for nearly all risk premia. As Merton (1973) emphasize, variation across assets in expected returns is driven by variation across assets in covariances with *recessions* far more than by variation across assets in covariances with consumption growth. Investors fear stocks because they do badly in occasional serious recessions unrelated to the risks of long-run average consumption growth.

Conclusion

The model is calibrated to fit the unconditional equity premium and riskfree interest rate. The model then generates long-horizon predictability of excess stock and bond returns from the dividend-price ratio, and mean reversion in returns; it generates high stock prices and return volatility despite smooth and unpredictable dividend streams; and it generates persistent movements in return volatility. The model predicts many puzzles that face the standard power utility consumption-based model, including the equity premium and riskfree rate puzzles and the low unconditional correlation of consumption growth with stock returns. The model is consistent with an even sharper long-run equity premium puzzle that results from mean-reversion in stock prices, together with low long-run consumption volatility.
The model supports the following view of risk premia in asset markets: *Individuals fear stocks primarily because they do badly in recessions (times of low surplus con-*

sumption ratios), not because stock returns are correlated with declines in wealth or consumption.

The parameter values in the calibrated model imply that habits are only about 5 percent lower than consumption on average. This degree of habit formation may seem rather extreme. However standard devices to increase the equity premium are not used. Some of them are concentration of stock ownership on a subset of the population or occasionally extremely bad states in the consumption distribution.

7 The Habit Formation Model and German Data

7.1 Results for German Data

I use data described in chapter five to calibrate the model. Since the excess return is much lower in German data γ will be much lower. For US data Campbell and Cochrane picked $\gamma = 2$ to match the Sharpe ratio[1]. In my calibration plausible values for γ are between 0.3125 and 0.3625.
The parameter γ looses its interpretation as a measure of risk aversion for lotteries over consumption, because utility is not just defined over consumption. Utility is defined over states and the agents have preferences over states. The state variable is the consumption surplus ratio. The state variable S is a function of consumption and habit. It is always positive since habit is always below consumption. The question is now if agents preferences over the constructed states are economically plausible. I will try to answer this question in chapter eight.
Calibrating the model at the sample average $\frac{1}{T}\sum_{t=1}^{T}\frac{P}{D}$ makes it possible to evaluate the performance of the model, because the model predicts the correct P/D-ratio by construction. Since the average is predicted correctly the deviations of P/D-ratios from the mean should explain the variation in real data.
Calibration one uses $\gamma = 0.3125$ to minimize the prediction error when consumption data are fed to the model and it also calibrates the model to match the Sharpe ratio of about 0.17 for the claim to consumption. I took the Sharpe ratio for the consumption claim in the GAUSS-output of the original program code[2]. So the first calibration would be analogous to Campbell and Cochrane. The results are reported completely for this calibration. Calibration two uses $\gamma = 0.3625$ to make the dividend claim in the model predict the average price-dividend ratio in the data correctly. Calibration three uses $\gamma = 2$ like Campbell and Cochrane did. I report the results for this calibration very briefly, but extensively enough to show that this calibration

[1] It is unclear how they got $\gamma = 2$. The equations 6.12, 6.14 and 6.15 should be solved simultaneously to find γ. The program `findgamCorrect.m` does this and the resulting γ is different from the one reported by Campbell and Cochrane. Further it is unclear if an annual value of $\sigma = 0.015$ or of $\sigma = 0.0122$ is chosen for the calibration. These different values appear in tables 6.1 and 6.2 and of course also in the original paper. It seems that they used $\sigma = 0.0122$ in the GAUSS-code, `main3.prg`.

[2] file: `results for gamma03125.txt`

is inappropriate for the German data. All other parameters remain unchanged for the different calibrations. I use the original GAUSS-code provided on the homepage of Cochrane[3]. The parameters in the code are changed according to German data. I use a finer grid for the state space. I have made changes only in the file `main3.prg` which I renamed to `main3d.prg`[4].

[3]gsb.uchicago.edu/fac/john.cochrane

[4]The GAUSS-outputs are stored in the folder `GAUSSOUT`.

7.1.1 Calibration One

The risk aversion parameter γ is chosen to minimize the prediction error for the price-dividend ratio when consumption data are fed to the model. The Sharpe ratio for the consumption claim is the same as in the data. Table 7.1 summarizes the parameter values. The state space is smaller as compared to US data.

Assumed Parameters

Mean consumption growth* g,%	2.427
Standard deviation of consumption growth* σ,%	1.743
Log riskfree rate* r^f,%	1.345
Persistence coefficient* ϕ	0.784
Utility curvature γ	0.3125
Standard deviation of dividend growth* σ_w,%	7.941
Correlation between Δd and Δc, ρ	0.094

Implied Parameters

Subjective discount factor* δ	0.9575
Steady state surplus consumption ratio $\bar{S}$	0.020
Maximum surplus consumption ratio S_{max}	0.032

Starred (*) values are annualized values, e.g. 12g, $\sqrt{12}\sigma$, $12r^f$, ϕ^{12}, δ^{12} since the model is simulated at a monthly frequency.

Table 7.1: Parameter choices

Table 7.2 shows expectations and standard deviations of historical and artificial data. As mentioned before the Sharpe ratios for historical data and artificial data of the consumption claim are identical. This was the target of this calibration. Excess returns are higher in artificial data. The standard deviation of excess returns is also too high and so are the price-dividend and price-consumption ratios.
Figure 7.1 shows the density of s for calibration one. The probability mass is much more concentrated at s_{max}. The fat tail is still there, but the bad states are less likely as compared to the calibration of Campbell and Cochrane. The area under the curve is clearly not equal to one so the scaling of the ordinate is wrong. This was also the case in the original paper.

Statistic	Consumption claim	Dividend claim	German data
$E(\Delta c)$	2.427*		2.427
$\sigma(\Delta c)$	1.743*		1.743
$E\left(r^f\right)$	1.34*		1.34
$E\left(r-r^f\right)/\sigma\left(r-r^f\right)$	0.17*	0.13	0.17
$E\left(r-r^f\right)$	3.87	3.74	2.714
$\sigma\left(r-r^f\right)$	23.44	28.40	14.099
$exp\left[E(p-d)\right]$	40.60	42.62	34.724
$E(P/D)$	42.15	44.35	35.98
$\sigma(p-d)$	0.33	0.35	0.269

Table 7.2: Means and standard deviations of data

The model is simulated at a monthly frequency. Statistics are calculated from artificial data at an annual frequency. Asterisks denote statistics that model parameters were chosen to replicate. All returns are annual percentages, $\gamma = 0.3125$.

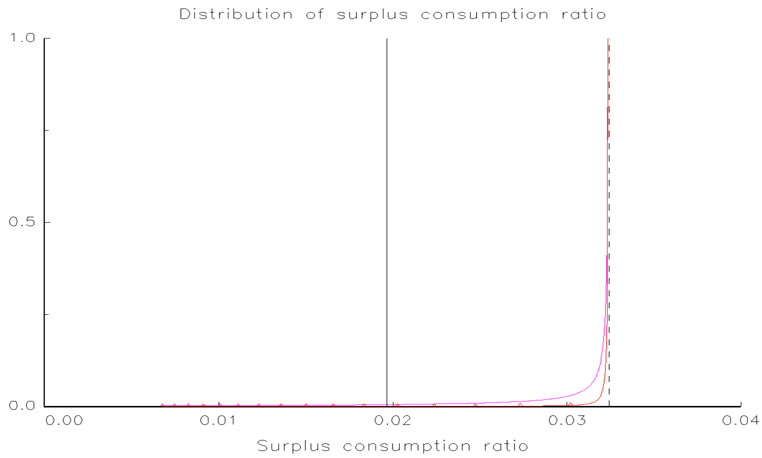

Figure 7.1: Density of s, $\gamma = 0.3125$

Figure 7.2 shows the sensitivity function $\lambda(s)$. The sensitivity function provides heteroskedasticity for the first order autoregressive process. In this calibration $\lambda(s)$ is higher for all states compared to the calibration for the U.S. data. So the heteroskedasticity as well as the variance will be higher.
Figure 7.3 plots the sensitivity of habit to consumption. It shows the quotient of changes of $\ln(X) = x$ and $\ln(C) = c$ so the graphic is hard to interpret. Consumption and habit always move in the same direction since dx/dc is always positive. At the steady state changes of consumption do not affect habit. In the two extreme states $S = 0$ and $S = S_{max}$ habit moves one to one with consumption. At S_{max} the consumption surplus ratio must be prevented from rising over S_{max} and at $S = 0$ from falling below zero.

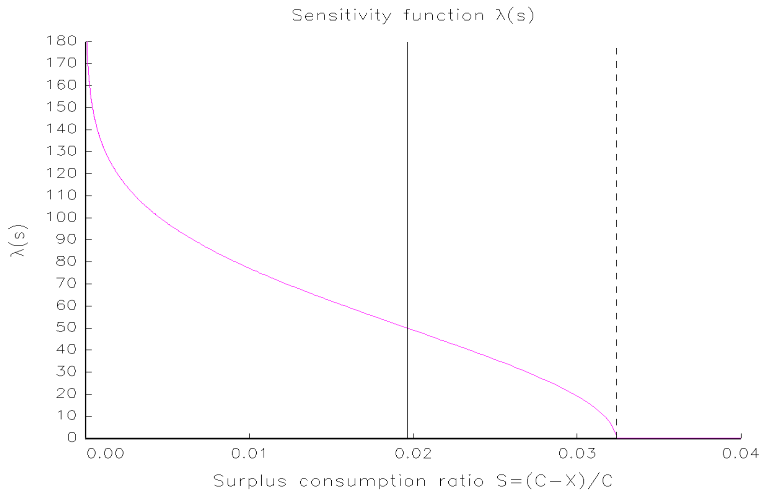

Figure 7.2: Sensitivity function λ(s)

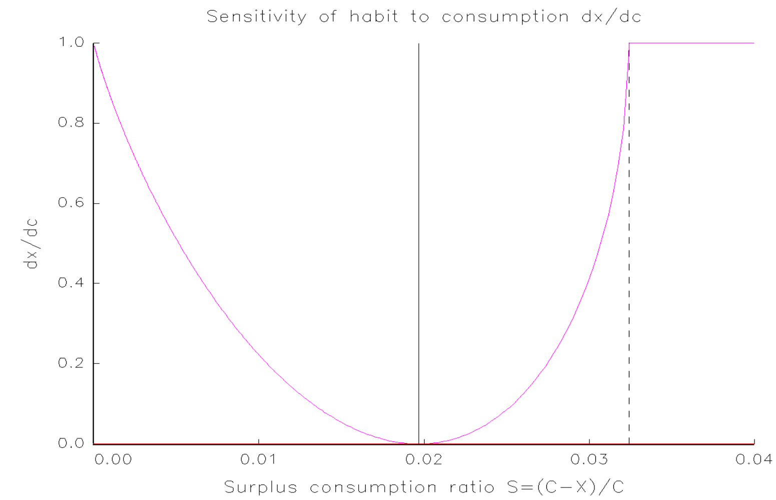

Figure 7.3: Sensitivity of habit to consumption

Figure 7.4 shows the price-consumption and price-dividend ratios as functions of the state variable s. In "good" states i.e. states with $S > \bar{S}$ agents pay higher prices for both claims. Campbell and Cochrane refer to those states as boom. In "bad" states with $S < \bar{S}$ agents pay less for both claims. Campbell and Cochrane name those states recessions. The price dividend and price consumption ratios are higher for the calibration to German data. As I mentioned before the model is calibrated to predict the average price dividend ratio correctly. The price consumption and price dividend ratios are limited to $S = 0$ and $S = S_{max}$ so that the model cannot generate values beyond $\frac{P}{C}(S = 0)$, $\frac{P}{C}(S_{max})$, $\frac{P}{D}(S = 0)$ and $\frac{P}{D}(S_{max})$ although those values are observed in real data. So the model is unable to generate certain values by construction, because the consumption surplus ratio is unable to take negative values or to exceed a certain level of S_{max}. There may be some justification for a

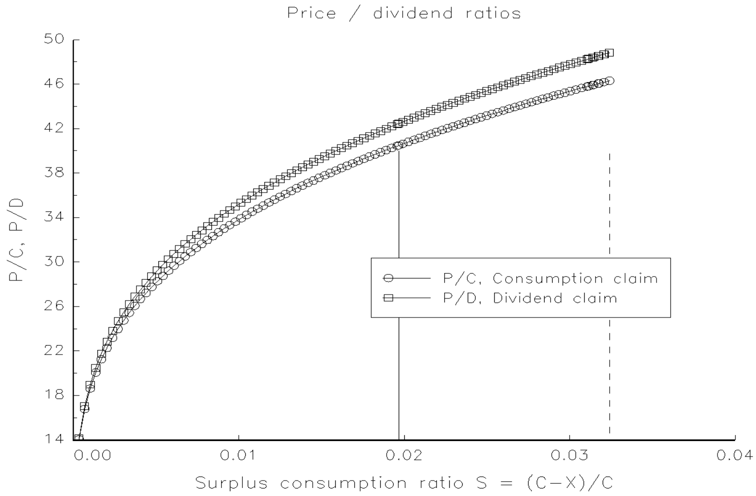

Figure 7.4: Price-dividend ratios

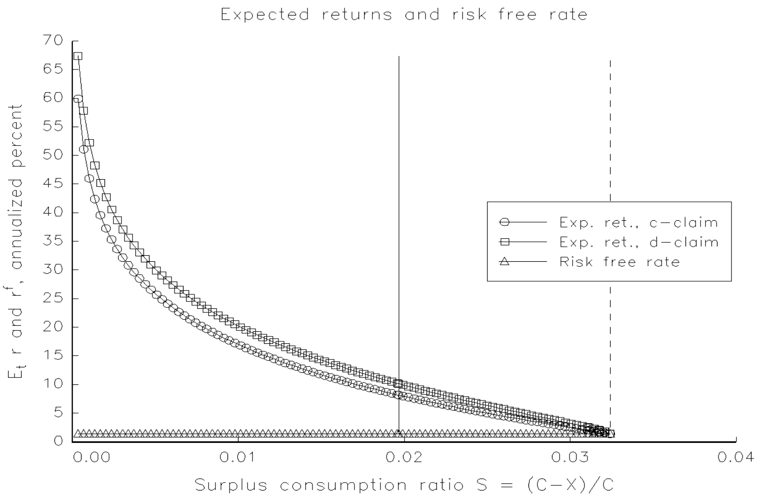

Figure 7.5: Expected returns

positive value of S, but there is no economic reason to stop S at S_{max}.
Figure 7.5 plots the expected return against the consumption surplus ratio. The riskfree rate is constant by construction. The dividend claim has a more highly expected return than the claim to consumption because of the extra risk. At $S = 0$ the expected returns of the two risky assets go up to infinity and at S_{max} all assets are expected to pay nearly the same returns. This leads to an interesting interpretation of that state. At S_{max} the price of risk is close to zero, but the curvature at S_{max} is $\eta = \gamma/S_{max} = 0.3125/0.032 = 9.766$. All three assets are expected to pay higher returns as compared to the original calibration of Campbell and Cochrane.

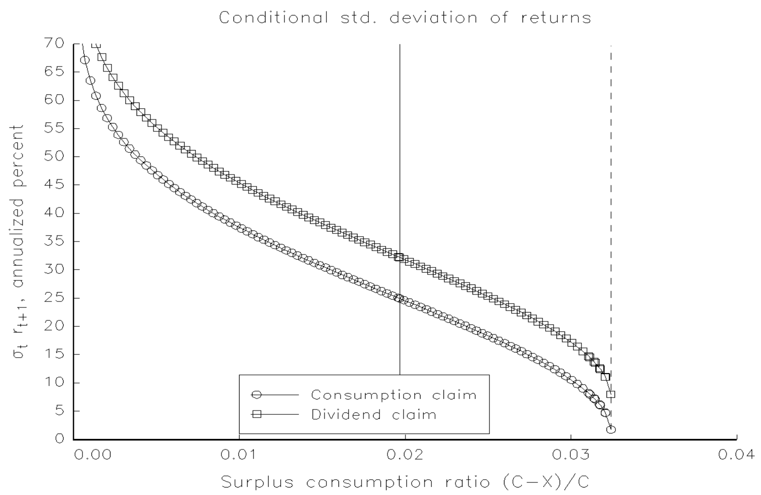

Figure 7.6: Conditional standard deviations of returns

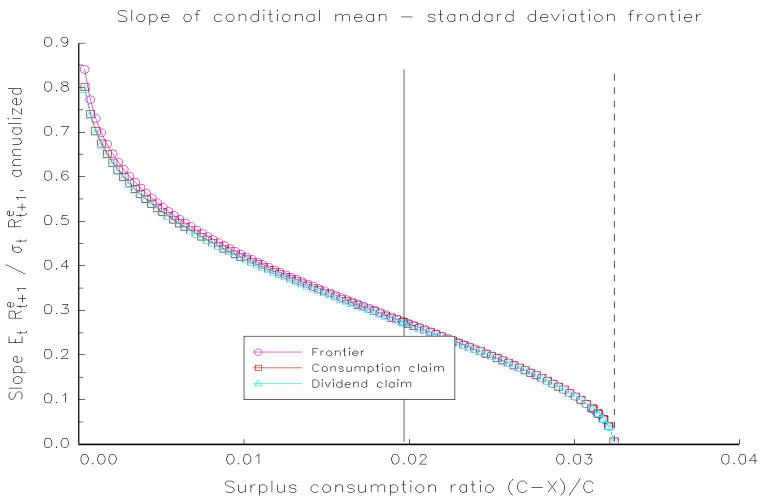

Figure 7.7: Sharpe ratios as functions of the surplus consumption ratio

The conditional standard deviations are shown in Figure 7.6. They go up to infinity as S goes to zero and down to $\sigma = 1.743$ and $\sigma_w = 7.941$ at S_{max}, because $\lambda\,(s_{max})$ is zero and so the only sources of volatility are σ and σ_w. The extra risk of the dividend claim is not fully priced and so does not show up in the expected returns because only part of the risk is correlated with consumption.
Figure 7.7 shows the Sharpe ratio. Both assets are nearly mean-variance efficient and lie close to the mean-variance frontier. Standard deviations are higher as compared to the original calibration. The expected standard deviation of the dividend claim is lower as compared to the original calibration for states close to S_{max}.
Table 7.3 shows autocorrelation of the data. The first lag is equal to $\phi = 0.78$ by construction, since the price-dividend and price-consumption ratios are functions of s. The second lag is a perfect match. For the remaining lags the model generates too low autocorrelations. For the excess returns the model generates small negative

autocorrelations like in the original paper of Campbell and Cochrane (Fama and French 1988a). So price changes tend to be reversed. All lags are predicted with the right sign except for lag three. The partial sum of autocorrelation coefficients is negative for all lags in artificial and historical data. All values are predicted with the right sign, but the values in historical data are more negative. The autocorrelation of absolute returns shows no uniform pattern. This statistical characteristic is to indicate heteroskedasticity. Since I have used annual data, this might not be the case. It is well known that heteroskedasticity increases with time frequency in return data.

		Lag				
Variable	Source	1	2	3	5	7
$p-d$	Cons. claim	0.78	0.61	0.48	0.31	0.19
	Dividend claim	0.78	0.61	0.48	0.31	0.19
	German data	0.78	0.61	0.54	0.34	0.28
$r-r^f$	Cons. claim	-0.09	-0.08	-0.06	-0.02	-0.01
	Dividend claim	-0.09	-0.08	-0.06	-0.02	-0.01
	German data	-0.16	-0.37	0.14	-0.11	-0.07
$\sum_{i=1}^{j} \rho\left(r_t^e, r_{t-j}^e\right)$	Cons. claim	-0.09	-0.17	-0.23	-0.28	-0.33
	Dividend claim	-0.09	-0.18	-0.23	-0.29	-0.34
	German data	-0.16	-0.53	-0.39	-0.39	-0.67
$\lvert r \rvert$	Cons. claim	0.40	0.31	0.25	0.15	0.09
	Dividend claim	0.36	0.28	0.22	0.13	0.08
	German data	-0.22	0.33	0.02	-0.00	-0.24

Table 7.3: Autocorrelations of data

The model values are based on time-aggregated values using a monthly simulation interval. Data are all annual. The third block of numbers is a partial sum of return autocorrelations until lag j.

Table 7.4 presents cross-correlations of data. In artificial data high prices imply low returns for all future horizons. For the German data the third and seventh lag are positive. So for two years ahead high prices imply low returns. The cross-correlations of excess return, logarithmic price-dividend and price-consumption ratios with absolute excess returns are ment to measure the leverage effect (Black 1976). There is no uniform picture. Negative returns cause subsequent high volatility and high returns imply low volatility. High price-dividend ratios also imply subsequent high volatility. The model data contain the leverage effect.

Table 7.5 presents long-horizon regressions of log excess stock returns on the log price-dividend ratio in artificial and historical data. All coefficients are negative: high prices imply low expected returns (Fama and French 1988b, Campbell and

Shiller 1988b). As in the original paper the R^2 starts low but then increases in artificial data. The R^2 is higher for the consumption-claim than for the dividend-claim due to the extra noise. In real data the R^2 also rises up to a five year horizon. Unfortunately, Campbell and Cochrane do not report t-statistics. For the real data all parameters are insignificant at a five percent level. All parameters have the right sign and decrease in artificial data as well as in real data.

		Lag				
Variable	Source	1	2	3	5	7
$p_t - d_t, r^e_{t+j}$	Cons. claim	-0.43	-0.32	-0.24	-0.15	-0.10
	Dividend claim	-0.37	-0.28	-0.21	-0.13	-0.08
	German data	-0.27	-0.08	0.03	-0.09	0.19
$r^e_t, \lvert r^e_{t+j}\rvert$	Cons. claim	-0.18	-0.13	-0.09	-0.06	-0.03
	Dividend claim	-0.20	-0.15	-0.11	-0.06	-0.03
	German data	-0.10	0.13	0.07	0.00	0.01
$p_t - d_t, \lvert r^e_{t+j}\rvert$	Cons. claim	-0.71	-0.54	-0.42	-0.26	-0.16
	Dividend claim	-0.65	-0.50	-0.38	-0.24	-0.15
	German data	0.14	0.30	0.16	0.05	0.23

Table 7.4: Cross-correlations of data

Horizon	C. claim		D. claim		German data	
(Years)	10×coef.	R^2	10×coef.	R^2	10×coef.	R^2
1	-3.00	0.19	-3.03	0.14	-2.1	0.07
2	-5.28	0.32	-5.32	0.24	-3.1	0.07
3	-6.99	0.41	-7.03	0.31	-4.4	0.11
5	-9.34	0.53	-9.39	0.40	-7.3	0.16
7	-10.91	0.60	-10.97	0.45	-6.1	0.09

Table 7.5: Long-horizon return regressions

In the model, consumption growth and consumption claim returns are conditionally perfectly correlated, since consumption growth is the only source of uncertainty. The error term is the same for the consumption random walk and for the AR(1) process that generates the states. The relation between consumption growth and return varies over time with the surplus consumption ratio S. Lambda is a function of s and varies the effect of the error term on the state in the next period. So for monthly data all datapoints are between two straight lines and each line corresponds to one limit of the consumption surplus ratio[5]. Figure 7.8 plots simulated monthly

[5] If λ would be constant all points would lie on a line.

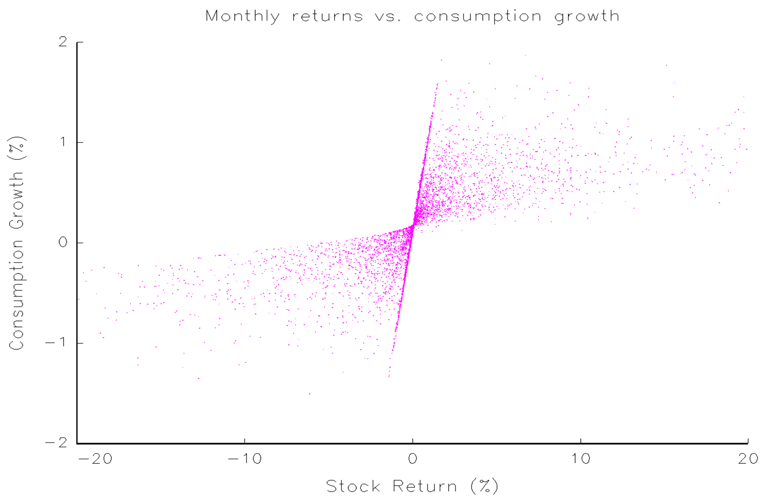

Figure 7.8: Simulated monthly consumption growth vs. monthly consumption claim returns

consumption growth against monthly consumption claim returns. The pattern of the scatterplot makes clear that consumption growth and consumption claim returns are positively correlated and move strikly positively together since all observations are bounded by the two lines. Time-aggregation changes the pattern. On a yearly basis consumption and returns can move more independently as figure 7.9 shows.

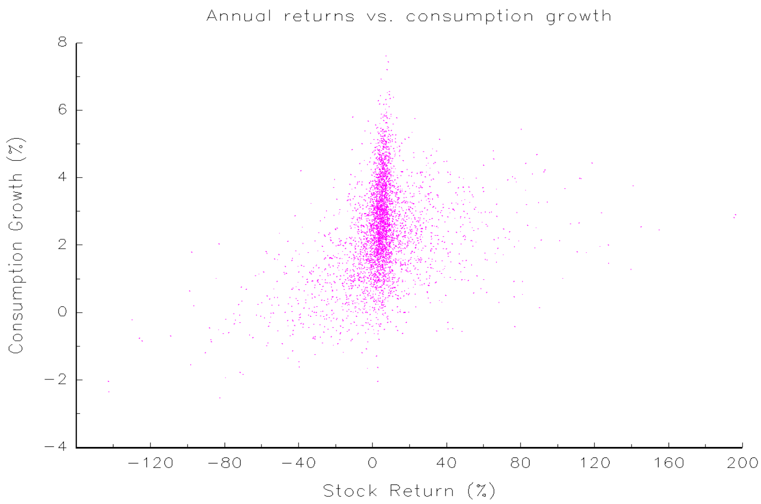

Figure 7.9: Growth in simulated annual consumption vs. annual consumption claim returns

Table 7.6 presents correlation between consumption growth and stock returns for different leads and lags. For monthly data there is only a simultaneous correlation. There is no correlation for leads or lags. For annual data the correlation is negative for lagged the consumption growth but positive for simultaneous consumption and the simultaneous correlation decreases with the time aggregation. Correlation is also

positive between consumption in the following year and excess returns. German data show a similar picture. The lagged consumption is also negative. The simultaneous correlation is negative contrary to the model data. The future consumption is positively correlated with the stock returns. An explanation would be the ability of financial markets to anticipate a period of stronger consumption growth. The model fails to generate positive correlation for more leads. It shows nearly no correlation in this calibration and not in the original calibration either.

	monthly		annual		
Corr.	Cons. claim	Div. claim	Cons. claim	Div. claim	Germany ann.
$r_t^e, \Delta c_{t-2}$	0.0	0.0	-0.13	-0.11	-0.12
$r_t^e, \Delta c_{t-1}$	0.0	0.0	-0.16	-0.14	-0.41
$r_t^e, \Delta c_t$	0.64	0.76	0.28	0.39	-0.25
$r_t^e, \Delta c_{t+1}$	0.0	0.0	0.34	0.41	0.15
$r_t^e, \Delta c_{t+2}$	0.0	0.0	0.0	0.0	0.29

Table 7.6: Correlation between consumption growth and stock return

	Correlation of the discount factor with		
	Consumption growth	Consumption claim return	Dividend claim return
Monthly	0.60	0.95	0.94
Monthly	0.27	0.95	0.92

Table 7.7: Correlation between the discount factor $M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \frac{S_{t+1}}{S_t} \right)^{-\gamma}$ and consumption growth, consumption claim return and dividend claim return.

Table 7.7 presents the correlation of the stochastic discount factor with consumption growth, dividend claim return and consumption claim return. Discount factor proxies which are better correlated with the discount factor produce smaller pricing errors for a given set of assets. The logarithm of the discount factor can be written as

$$m_{t+1} = \ln(\delta) - \gamma \left(\Delta c_{t+1} + \Delta s_{t+1} \right). \tag{7.1}$$

The discount factor is driven by the growth rates of consumption and the consumption surplus ratio. The two growth rates are stochastic processes which are driven by the error term v_{t+1}. To understand the movement of the discount factor one needs to know the state variable s and the realisation of the error term. So all variables correlated with s and especially all variables which are functions of s will be more correlated with the discount factor. In artificial monthly data and especially yearly

data the returns on dividend claim and consumption claim are better correlated with m than consumption growth[6]. The discount factor is conditionally perfect correlated with consumption growth[7], the unconditional correlation is low because of the variation in the consumption surplus ratio.
Now historical data are fed to the model. Assuming that the consumption surplus ratio starts at $\bar{s}$ figure 7.10 displays consumption and habit. Habit is always below consumption. After the oil crisis and after the unification of Germany habit comes very close to consumption. Cyclical dips in consumption growth bring habit closer to consumption. The model predicts low price-dividend ratios for these years.
Figure 7.11 compares historical price-dividend ratios and price-dividend ratios implied by the model. There is some comovement of artificial data and historical data. Like in the US data the worst performance of the model happens for the years after about 1990. The model curve moves downwards and price-dividend ratios in historical data move upwards for the German data as well as for the US data.

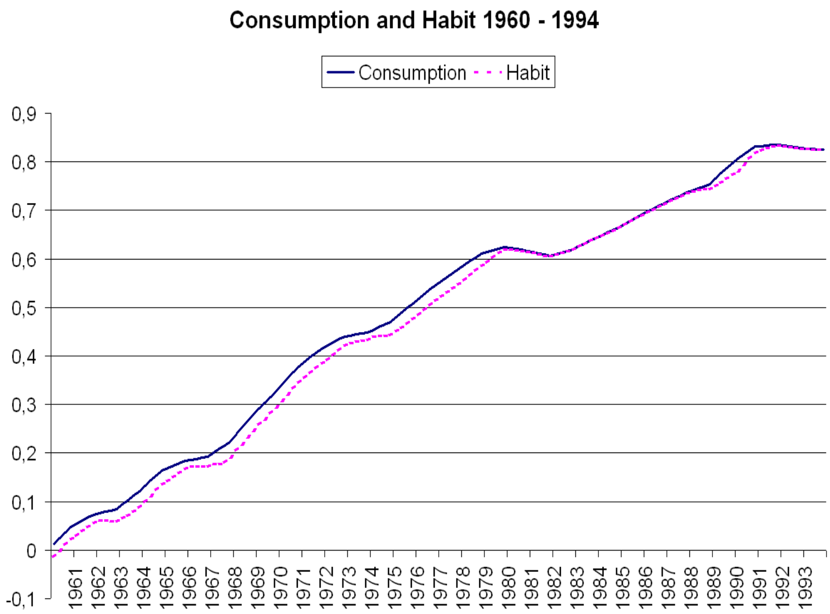

Figure 7.10: Nondurables and services consumption and habit level implied, assuming that s starts at $\bar{s}$

Table 7.8 summarizes some statistics of the data used to create Figure 7.11. The average price-dividend ratio is about 36 for model data and historical data[8]. The three performance measures show that the sample average of historical data fits the line better than the model data. Model data and historical data are negatively

[6]In the original paper, consumption growth is more correlated with m compared to the dividend claim return. This happens because the higher correlation of $\rho = 0.2$ adds more extra noise to the dividend claim compared to the correlation of 0.094.
[7]Conditioning on S gives a unique solution for m if one knows the error term.
[8]This was the target of the calibration using $\gamma = 0.3125$ as well as matching the Sharpe ratio

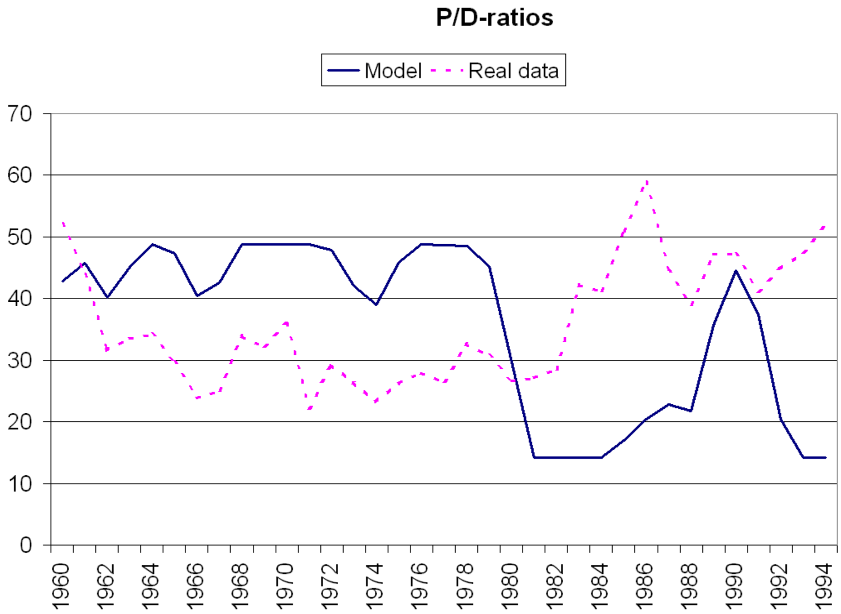

Figure 7.11: Historical price-dividend ratio, and model prediction based on the history of consumption

correlated. The log growth rates are positively uncorrelated. So at least the model curve and the curve made from historical data have some movement in the same direction and can predict the direction of change of the real data.

Mean values

	hist. data	model data
$\frac{1}{T}\sum_{t=1}^{T}\frac{P_t}{D_t}$	36	36

Measures to evaluate the performance

$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{P_t^M}{D_t^M}-\frac{P_t^H}{D_t^H}\right)^2}$	20
$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\frac{P_t^H}{D_t^H}-\frac{1}{T}\sum_{t=1}^{T}\frac{P_t}{D_t}\right)^2}$	10
$\rho\left(\frac{P_t^M}{D_t^M},\frac{P_t^H}{D_t^H}\right)$ $\rho\left(\Delta\left(p-d\right)_{t+1}^{M},\Delta\left(p-d\right)_{t+1}^{H}\right)$	-0.48 0.214

Table 7.8: Statistics of data used in Figure 7.11

7.1.2 Calibration Two

The target of this calibration is to match the average price-dividend ratio of the model with the real data. Since all figures look quite similar to calibration one I will not present them. I will present only interesting tables. Table 7.9 shows assumed and implied parameters for this calibration. The state space has grown since s_{max} grows with gamma. Table 7.10 shows the statistics already known from calibration one. Risk aversion has increased, so price-consumption and price-dividend ratios are expected to be lower and expected excess returns are expected to increase. The average price-consumption ratio has fallen from 42.15 to 35.30 and the average price-consumption ratio from 44.35 to 35.97. The Sharpe ratio for the consumption claim has increased from 0.17 to 0.18 and for the dividend claim from 0.13 to 0.15. Excess returns go up from 3.87 to 4.41 and from 3.74 to 4.38, so these statistics move even more away from real data as compared to calibration one.

Assumed Parameters

Mean consumption growth* g,%	2.427
Standard deviation of consumption growth* σ,%	1.743
Log riskfree rate* r^f,%	1.345
Persistence coefficient* ϕ	0.784
Utility curvature γ	0.3625
Standard deviation of dividend growth* σ_w,%	7.941
Correlation between Δd and Δc, ρ	0.094

Implied Parameters

Subjective discount factor* δ	0.953
Steady state surplus consumption ratio $\bar{S}$	0.021
Maximum surplus consumption ratio S_{max}	0.035

Table 7.9: Parameter choices

Statistic	C. claim	D. claim	German data
$E(\Delta c)$	2.427*		2.427
$\sigma(\Delta c)$	1.743*		1.743
$E(r^f)$	1.34*		1.34
$E(r-r^f)/\sigma(r-r^f)$	0.18*	0.15	0.17
$E(r-r^f)$	4.41	4.38	2.714
$\sigma(r-r^f)$	24.02	29.24	14.099
$exp[E(p-d)]$	33.91	34.44	34.724
$E(P/D)$	35.30	35.97	35.98
$\sigma(p-d)$	0.34	0.36	0.269

Table 7.10: Means and standard deviations of data

7.1.3 Calibration Three

This calibration is just presented to show that $\gamma = 2$ is not an appropriate way to calibrate the model. In figures 7.11 and 7.12 all statistics except for the standard deviations of excess returns and price-dividend ratios have moved away even more from the statistics made from real data.

7.1.4 Which Calibration is to be Choosen?

Calibration one matches the Sharpe ratio and generates a very close match for the sample averages of model data and real data if consumption is fed to the model. It produces the smallest expected excess returns which come closest to the real data.

Assumed Parameters

Mean consumption growth* g,%	2.427
Standard deviation of consumption growth* σ,%	1.743
Log riskfree rate* r^f,%	1.345
Persistence coefficient* ϕ	0.784
Utility curvature γ	2.000
Standard deviation of dividend growth* σ_w,%	7.941
Correlation between Δd and Δc, ρ	0.094

Implied Parameters

Subjective discount factor* δ	0.8141
Steady state surplus consumption ratio $\bar{S}$	0.050
Maximum surplus consumption ratio S_{max}	0.082

Table 7.11: Parameter choices

Statistic	C. claim	D. claim	German data
$E(\Delta c)$	2.427*		2.427
$\sigma(\Delta c)$	1.743*		1.743
$E\left(r^f\right)$	1.34*		1.34
$E\left(r - r^f\right)/\sigma\left(r - r^f\right)$	0.62*	0.57	0.17
$E\left(r - r^f\right)$	9.24	12.46	2.714
$\sigma\left(r - r^f\right)$	14.81	21.89	14.099
$exp\left[E(p-d)\right]$	12.48	8.94	34.724
$E(P/D)$	12.71	9.14	35.98
$\sigma(p-d)$	0.20	0.22	0.269

Table 7.12: Means and standard deviations of data

So this calibration is best. Standard deviations are too high and become better for higher values of γ. I think expected returns and standard deviation of expected returns are more interesting quantities than price-dividend and price-consumption ratios. So the Sharpe ratio would be the central quantity since it measures the slope of the mean-variance frontier and is a measure of risk. The payment of dividends depends on the economic situation but also on many other factors like taxation rules and the way corporate finance is done. If interest rates and stock prices are high and the company makes good profits like in a "boom", it may be cheaper to keep the money inside the company to finance investments. In this case the state variable S would be greater than $\bar{S}$. If interest rates and stock prices are low and the company has little investment opportunities like in a "recession", shareholders may force the company to get rid of cash by paying dividends. In this case S would be below $\bar{S}$. This would be a scenario for a business cycle like in the paper. But interest rates,

taxation and especially changes in taxation will also be important. So price-dividend ratios remain a noisy indicator and are hard to predict.
The mismatch in standard deviations of calibration one is due to the shape of the distribution of the state variable s and the change in the functions of the price-dividend ratios and excess returns. The distribution is much more left skewed than in the calibration for the US market. As a result the standard deviations of all statistics increase since all variables are functions of s. The construction of a state process that keeps its skewness and standard deviation relative to the interval of the state space for different values of γ could solve the problem. Table 7.13 shows the standard deviations of S for calibrations one and three. S moves stronger along the state space in calibration one.

Statistic	Standard-deviation	S_{max}	Stddev/S_{max}
Calibration one	0.0104	0.032	0.325
Calibration three	0.0189	0.082	0.230

Table 7.13: Standard deviations of S

8 A Critical Review of the Habit Model

In this part I would like to discuss the construction of the model. In my opinion the central equation is the state generating AR(1) process for s. Since all other quantities are functions of the state variable, the process must be constructed in a way so that all other variables, which are functions of s, have statistical properties like in the real data. The first order autocorrelation ϕ is used to calibrate the process. So the first order autocorrelation of price-dividend ratio and price-consumption ratio is a perfect match by construction. Other exogenous parameters are γ and σ which enter the process over $\bar{s}$ and $\lambda(s)$. This function provides heteroskedasticity and controls the upper bound. All functions of s e.g. expected excess returns and price-dividend ratios are limited at s_{max}. If there are for example higher values of price-dividend ratios than $\frac{P}{D}(s_{max})$, the model is unable to predict them by construction.
One could estimate the time series characteristics of price-dividend ratios and excess returns for real data. Since price-dividend ratios and Sharpe ratios are functions of s, a cleverly designed exogen process could generate values for s so that the statistical properties of price-dividend ratios and Sharpe ratios show the same pattern as in the real data. The log price-dividend and log price-consumption ratios show the same autocorrelation pattern as s. So the construction of the process gives all structure and statistical properties to the depending variables and this process is exogenous. A riskfree rate puzzle is avoided, because the riskfree rate is constant by construction. The central point in the consumption-based model is the high risk aversion expressed by a very high value of γ. This parameter is too high if one interprets the CRRA utility function as preferences over lotteries of consumption. By calculation of the certainty equivalent one can answer the question if a certain degree of risk aversion is reasonable.
Some open questions remain. Why is this process constructed in this way? Why is the sensitivity function constructed in such a way and what is its justification?

8.1 On the Interpretation of Risk Aversion

Usually the CRRA utility function is used to describe utility over uncertain outcomes.

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma} \tag{8.1}$$

The Arrow-Pratt coefficient of relative risk aversion is defined as

$$r_R(c) = \frac{-u''(c)}{u'(c)}. \tag{8.2}$$

For the CRRA utility function $r_R(c)$ is equal to γ. The Habit formation model uses curvature given in equation (6.3). To decide if a certain degree of risk aversion measured by the Arrow-Pratt coefficient of relative risk aversion is reasonable, one has to consider the *certainty equivalent*.

Definition 8.1 (certainty equivalent) *Given a utility function u(·) the certainty equivalent of F(·), denoted c(F,u) is the amount of money for which the individual is indifferent between the lottery F(·) and the certain amount c(F,u); that is,*

$$u(c(F,u)) = \int u(x)\,dF(x)$$

(Mas-Colell, Whinston and Greene 1995)

Using the calibration of Campbell and Cochrane for postwar data $\bar{S} = 0.057$ and $\gamma = 2$ and so $\eta = \gamma/\bar{S} = 35$ which is a high degree of risk aversion at the steady state. Since $S_{max} = 0.094$ and $S_{min} > 0$ the Arrow-Pratt coefficient $\eta \in [21, \infty)$ can explode. In bad states risk aversion goes up to infinity.
In the standard consumption-based model the riskfree rate is too high as a result of the risk aversion parameter in the CRRA utility function. In the habit formation model the riskfree rate is constant, but there is still the certainty equivalent left to decide if the risk aversion is reasonable. Utility is now defined on states and an interpretation is hard because of habit. Nevertheless, I will report them in the next table[1]. The division of riskpremia through the size of the state space shows that the riskpremia increase relative to S_{max} for higher values of γ. All I can say is that for example in calibration one an agent would value a lottery over states like in calibration one as high as a state value of 0.0253 although the average state value of the distribution would be 0.0264. The quotient X/C shows which fraction of

[1] All calculations are done using `CertEqu2.m`.

consumption is habit. If agents are in urgent need of this large part of consumption, the degree of habit formation is extreme. The normal citizen in the United States or Germany would be in urgent need of about 94% of his spendings on nondurable goods and services.
Maybe there are other ways of interpreting habit in this context. It would be nice to interpret it as a measureable variable so that empirical tests could be made.

Calibration	$E(S)$	$E(X/C)$	Certainty-equivalent	$(X/C)_{CE}$	Risk-premia	RP/S_{max}
Campbell	0.0623	0.9377	0.0481	0.9519	0.0140	0.1489
one	0.0264	0.9736	0.0253	0.9747	0.0011	0.0343
two	0.0278	0.9722	0.0265	0.9735	0.0014	0.0400
three	0.0547	0.9453	0.0422	0.9578	0.0124	0.1512

Table 8.1: Riskpremia for different calibrations

9 Conclusion

The task of this thesis was to calibrate the habit formation model of Campbell and Cochrane to German data. I started with a general introduction to asset pricing using the stochastic discount factor approach. I presented the equity premium puzzle, the riskfree rate puzzle and the correlation puzzle using statistics of US data and German data. I discussed the original habit formation paper (Campbell and Cochrane 1994) and derived the central pricing equation. The calibration to German data was not done using a complete postwar data set, because suitable asset market data and consumption data were only available from 1960 to 1994. The excess return of shares over the riskfree rate is far lower than it would be for a full postwar data set starting from 1945. Nevertheless, the standard consumption model was unable to explain excess return and Sharpe ratio with a plausible amount of risk aversion in the CRRA utility function.

The calibration of the habit formation model to German data can also be regarded as a performance test in samples with lower excess return and Sharpe ratio. As a result the parameter of risk aversion γ was much lower as compared to the orginal calibration. I evaluated the model when real consumption data were fed into it and found some movement in the same direction of model data and historical data for the growth rates of price-dividend ratios in the German case. The absolute price-dividend ratios were negatively correlated. For US data I found a small negative correlation for the growth rates of price-dividend ratios, but a positive correlation for absolute price-dividend ratios.

In my opinion the complete structure in this model comes from the process of s. The real riskfree interest rate is constant by construction so there cannot be a riskfree rate puzzle. This was one way to show that a certain degree of risk aversion is unrealistic. Another way would be to think about risk premia as results from preferences over lotteries of consumption. In the habit formation model this approach would need measureability of habit or at least a proxy for habit. Since such a variable is hard to find, one cannot precisely assess the risk aversion in this model.

10 Appendix

10.1 Appendix A

$$\left|\frac{E\left(R^{mv}\right)-R^f}{\sigma\left(R^{mv}\right)}\right| = \frac{\sigma\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]}{E\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]} \tag{10.1}$$

since $\sigma^2(x) = E\left(x^2\right) - \left[E(x)\right]^2$,

$$\Rightarrow \sqrt{\frac{\sigma^2\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]}{\left\{E\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]\right\}^2}} = \sqrt{\frac{E\left\{\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]^2\right\}}{\left\{E\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]\right\}^2} - 1} \tag{10.2}$$

$$= \sqrt{\frac{E\left[\left(c_{t+1}/c_t\right)^{-2\gamma}\right]}{E\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]E\left[\left(c_{t+1}/c_t\right)^{-\gamma}\right]} - 1} \tag{10.3}$$

with lognormal consumption growth and using the fact that normal z means

$$E\left(e^z\right) = e^{E(z)+(1/2)\sigma^2(z)} \tag{10.4}$$

this is equal to

$$\sqrt{\frac{e^{-2\gamma E(\Delta \ln c_{t+1})+2\gamma^2\sigma^2(\Delta \ln c_{t+1})}}{e^{-2\gamma E(\Delta \ln c_{t+1})+2(\gamma^2/2)\sigma^2(\Delta \ln c_{t+1})}} - 1} \tag{10.5}$$

and so

$$\left|\frac{E\left(R^{mv}\right)-R^f}{\sigma\left(R^{mv}\right)}\right| = \sqrt{e^{\gamma^2\sigma^2(\Delta \ln c_{t+1})} - 1} \tag{10.6}$$

10.2 Appendix B

$$R^f = \frac{1}{\beta}\left(\frac{c_{t+1}}{c_t}\right)^{\gamma} \quad \Leftrightarrow \quad R^f = 1/\left\{E_t\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right]\right\} \tag{10.7}$$

Define,

$$r_t^f = \ln R_t^f \qquad \beta = e^{-\delta} \tag{10.8}$$

$$\Delta \ln c_{t+1} = \ln c_{t+1} - \ln c_t \tag{10.9}$$

Using the fact that normal z means

$$E\left(e^z\right) = e^{E(z)+(1/2)\sigma^2(z)} \tag{10.10}$$

we have

$$R_t^f = \left[e^{-\delta} e^{-\gamma E_t(\Delta \ln c_{t+1}) + \left(\gamma^2/2\right)\sigma_t^2(\Delta \ln c_{t+1})}\right]^{-1} \tag{10.11}$$

take the log

$$r_t^f = \delta + \gamma E_t\left(\Delta \ln c_{t+1}\right) - \frac{\gamma^2}{2}\sigma_t^2\left(\Delta \ln c_{t+1}\right) \tag{10.12}$$

10.3 Appendix C

Utility in period t is given by

$$u_t\left(C_t\right) = \frac{\left(C_t - X_t\right)^{1-\gamma} - 1}{1-\gamma} \tag{10.13}$$

The Arrow-Pratt coefficient of relative risk aversion is defines as

$$\eta_t \equiv -\frac{C_t u_{cc}\left(C_t, X_t\right)}{u_c\left(C_t, X_t\right)}. \tag{10.14}$$

$$\Leftrightarrow \eta_t = -\frac{C_t\left[-\gamma\left(C_t - X_t\right)^{-\gamma-1}\right]}{\left(C_t - X_t\right)^{-\gamma}} = \frac{C_t\gamma}{C_t - X_t} = \frac{\gamma}{S_t} \tag{10.15}$$

10.4 Appendix D

Take a loglinear approximation around the steady state $s = \bar{s}$ and $c_{t+1} - c_t = g$[1]. The first-order approximations are deraived as follows. The definition of s_t can be written

$$s_t = \ln\left(\frac{e^{c_t} - e^{x_t}}{e^{c_t}}\right). \tag{10.16}$$

The steady state of $x - c$ is derived as follows:

$$\begin{aligned} &\bar{S} = \frac{C - X}{C} \Leftrightarrow 1 - \bar{S} = 1 - \frac{C - X}{C} = 1 - \frac{C - C + X}{C} = \frac{X}{C} \\ \Longrightarrow\quad &\ln\left(1 - \bar{S}\right) = x - c = h \end{aligned} \tag{10.17}$$

Taking a linear approximation,

$$s_t - \bar{s} \approx \left(1 - \frac{1}{\bar{S}}\right)\left(x_t - c_t - h\right). \tag{10.18}$$

Similary,

$$\begin{aligned} \lambda\left(s_t\right)\left(c_{t+1} - c_t - g\right) &\approx \left[\lambda\left(\bar{s}\right) + \lambda'\left(\bar{s}\right)\left(s_t - \bar{s}\right)\right]\left(c_{t+1} - c_t - g\right) \\ &\approx \lambda\left(\bar{s}\right)\left(c_{t+1} - c_t - g\right). \end{aligned} \tag{10.19}$$

[1]This appendix is taken in part from Campbell and Cochrane 1995.

Then, equation (6.5) becomes:

$$
\begin{aligned}
\left(1-\frac{1}{\bar{S}}\right)(x_{t+1}-c_{t+1}-h)+\bar{s} &\approx \phi\left(1-\frac{1}{\bar{S}}\right)(x_t-c_t-h) \\
&\quad+\lambda(\bar{s})(c_{t+1}-c_t-g)+\bar{s} \\
\Longleftrightarrow \left(1-\frac{1}{\bar{S}}\right)(x_{t+1}-c_{t+1}-h) &\approx \phi\left(1-\frac{1}{\bar{S}}\right)(x_t-c_t-h)+\lambda(\bar{s})(c_{t+1}-c_t-g) \\
\Longleftrightarrow x_{t+1}-c_{t+1}-h &\approx \phi(x_t-c_t-h)+\frac{1/\bar{S}-1}{1-1/\bar{S}}(c_{t+1}-c_t-g) \\
\Longleftrightarrow x_{t+1}-c_{t+1}-h &\approx \phi(x_t-c_t-h)-(c_{t+1}-c_t-g) \\
\Longleftrightarrow x_{t+1} &\approx \phi(x_t-c_t-h)+c_t+g+h \\
&\approx \phi x_t-\phi c_t-\phi h+c_t+g+h \\
&\approx (1-\phi)c_t-(1-\phi)h-\phi x_t+g \\
&\approx [(1-\phi)h+g]+\phi x_t+(1-\phi)c_t
\end{aligned}
\tag{10.20}
$$

For x_t and x_{t-1} the equation becomes:

$$
\begin{aligned}
x_t &\approx [(1-\phi)h+g]+\phi x_{t-1}+(1-\phi)c_{t-1} \\
x_{t-1} &\approx [(1-\phi)h+g]+\phi x_{t-2}+(1-\phi)c_{t-2}
\end{aligned}
\tag{10.21}
$$

Now plug this in for x_t:

$$
\begin{aligned}
x_{t+1} &\approx [(1-\phi)h+g]+\phi\{[(1-\phi)h+g]+\phi x_{t-1}+(1-\phi)c_{t-1}\}+(1-\phi)c_t \\
&\approx [(1-\phi)h+g]+\phi[(1-\phi)h+g]+\phi^2 x_{t-1}+\phi(1-\phi)c_{t-1}+(1-\phi)c_t
\end{aligned}
\tag{10.22}
$$

Now plug in x_{t-1}:

$$
\begin{aligned}
x_{t+1} \approx\ & [(1-\phi)\,h+g]+\phi\,[(1-\phi)\,h+g] \\
& +\phi^2\left\{[(1-\phi)\,h+g]+\phi x_{t-2}+(1-\phi)\,c_{t-2}\right\} \\
& +\phi\,(1-\phi)\,c_{t-1}+(1-\phi)\,c_t \\
\approx\ & [(1-\phi)\,h+g]+\phi\,[(1-\phi)\,h+g] \\
& +\phi^2\,[(1-\phi)\,h+g]+\phi^3 x_{t-2}+\phi^2\,(1-\phi)\,c_{t-2} \\
& +\phi\,(1-\phi)\,c_{t-1}+(1-\phi)\,c_t \\
\approx\ & \left[\left(1+\phi+\phi^2\right)(1-\phi)\,h+\left(1+\phi+\phi^2\right)g\right]+\phi^3 x_{t-2}+\phi^2\,(1-\phi)\,c_{t-2} \\
& +\phi\,(1-\phi)\,c_{t-1}+(1-\phi)\,c_t \\
\approx\ & \left[\left(1-\phi^3\right)h+\left(1+\phi+\phi^2\right)g\right]+\phi^3 x_{t-2}+\phi^2\,(1-\phi)\,c_{t-2} \\
& +\phi\,(1-\phi)\,c_{t-1}+(1-\phi)\,c_t
\end{aligned}
\tag{10.23}
$$

So for infinite time and $\phi < 1$ this expression becomes:

$$
x_{t+1} \approx \left[h+\frac{g}{1-\phi}\right]+(1-\phi)\sum_{j=0}^{\infty}\phi^j c_{t-j} \tag{10.24}
$$

10.5 Appendix E

$$
M_{t+1} = \delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma} \tag{10.25}
$$

$$
\xrightarrow{\ln} m_{t+1} = \ln\delta-\gamma\left(s_{t+1}-s_t+c_{t+1}-c_t\right) \tag{10.26}
$$

Plug in the process $\Delta c_{t+1} = c_{t+1}-c_t = g+v_{t+1}$:

$$
\begin{aligned}
m_{t+1} &= \ln\delta-\gamma\left(g+s_{t+1}-s_t+v_{t+1}\right) \\
&= \ln\delta-\gamma g-\gamma\left(s_{t+1}-s_t+v_{t+1}\right)
\end{aligned}
\tag{10.27}
$$

$$
\xrightarrow{\ln^{-1}} M_{t+1} = \delta G^{-\gamma}e^{-\gamma\left(s_{t+1}-s_t+v_{t+1}\right)} \tag{10.28}
$$

Plug in (6.5) for s_{t+1}:

$$M_{t+1} = \delta G^{-\gamma} e^{-\gamma\{[(1-\phi)\bar{s}+\phi s_t+\lambda(s_t)(c_{t+1}-c_t-g)]-s_t+v_{t+1}\}} \tag{10.29}$$

Rearranging this equation and plugging in $\Delta c_{t+1} = g+v_{t+1} \Leftrightarrow -g = -c_{t+1}+c_t+v_{t+1}$:

$$M_{t+1} = \delta G^{-\gamma} e^{-\gamma[(1-\phi)(\bar{s}-s_t)+\lambda(s_t)v_{t+1}+v_{t+1}]} \tag{10.30}$$

$$M_{t+1} = \delta G^{-\gamma} e^{-\gamma[(\phi-1)(s_t-\bar{s})+(1+\lambda(s_t))v_{t+1}]} \tag{10.31}$$

10.6 Appendix F

$$\max_{\{all\ assets\}} \frac{E_t\left(R^e_{t+1}\right)}{\sigma_t\left(R^e_{t+1}\right)} = \frac{\sigma_t\left(M_{t+1}\right)}{E_t\left(M_{t+1}\right)} \tag{10.32}$$

$$\begin{aligned} \frac{\sigma(M)}{E(M)} &= \frac{\sqrt{E\left(M^2\right)-E\left(M\right)^2}}{E\left(M\right)} = \frac{\sqrt{e^{2\mu+2\sigma^2}-e^{2\mu+\sigma^2}}}{e^{\mu+\sigma^2/2}} \\ &= \sqrt{\frac{e^{2\mu+2\sigma^2}-e^{2\mu+\sigma^2}}{e^{2\mu+\sigma^2}}} = \sqrt{e^{\sigma^2}-1} \end{aligned} \tag{10.33}$$

Rewrite equation (10.31):

$$M_{t+1} = e^{\ln\delta-\gamma g-\gamma[(\phi-1)(s_t-\bar{s})+(1+\lambda(s_t))v_{t+1}]} \tag{10.34}$$

Let's get the variance of the exponent:

$$\begin{aligned} &Var\left\{\ln\delta-\gamma g-\gamma\left[(\phi-1)\left(s_t-\bar{s}\right)+\left(1+\lambda\left(s_t\right)\right)v_{t+1}\right]\right\} \\ &= \gamma^2 Var\left[\left(1+\lambda\left(s_t\right)\right)v_{t+1}\right] \\ &= \gamma^2\left[1+\lambda\left(s_t\right)\right]^2 Var\left[v_{t+1}\right] \\ &= \gamma^2\left[1+\lambda\left(s_t\right)\right]^2\sigma^2 \end{aligned} \tag{10.35}$$

Plug this in for σ^2 in (10.33).

10.7 Appendix G

$$R_t^f = \frac{1}{E_t\left(M_{t+1}\right)} \tag{10.36}$$

Take conditional expectation over M_{t+1}:

$$\left[E_t\left(M_{t+1}\right)\right]^{-1} = \left[E_t\left(e^{\ln\delta-\gamma g-\gamma[(\phi-1)(s_t-\bar{s})+(1+\lambda(s_t))v_{t+1}]}\right)\right]^{-1} \tag{10.37}$$

Take expectations using $E\left(e^z\right) = e^{E(z)+(1/2)\sigma^2(z)}$:

$$\left[E_t\left(M_{t+1}\right)\right]^{-1} = \left[e^{\ln\delta-\gamma g-\gamma(\phi-1)(s_t-\bar{s})+\left(\gamma^2\sigma^2/2\right)[1+\lambda(s_t)]^2}\right]^{-1} \tag{10.38}$$

$$R_t^f = e^{-\ln\delta+\gamma g-\gamma(1-\phi)(s_t-\bar{s})-\left(\gamma^2\sigma^2/2\right)[1+\lambda(s_t)]^2} \tag{10.39}$$

Take the log:

$$r_t^f = -\ln\left(\delta\right) + \gamma g - \gamma\left(1-\phi\right)\left(s_t-\bar{s}\right) - \frac{\gamma^2\sigma^2}{2}\left[1+\lambda\left(s_t\right)\right]^2. \tag{10.40}$$

10.8 Appendix H

Take equation (6.13) and plug in (6.15):

$$\begin{aligned} r_t^f =& -\ln\left(\delta\right) + \gamma g - \gamma\left(1-\phi\right)\left(s_t-\bar{s}\right) \\ & - \frac{\gamma^2\sigma^2}{2}\left[1+\left(1/\bar{S}\right)\sqrt{1-2\left(s_t-\bar{s}\right)}-1\right]^2 \end{aligned} \tag{10.41}$$

Rearrange:

$$\begin{aligned} r_t^f =& -\ln\left(\delta\right) + \gamma g - \gamma\left(1-\phi\right)\left(s_t-\bar{s}\right) \\ & - \frac{\gamma^2\sigma^2}{2}\frac{1}{\bar{S}^2}\left[1-2\left(s_t-\bar{s}\right)\right] \end{aligned} \tag{10.42}$$

Rearrange:

$$r_t^f = -\ln(\delta) + \gamma g - \left(\frac{\gamma}{\bar{S}}\right)^2 \frac{\sigma^2}{2}\left[\frac{2(1-\phi)(s_t-\bar{s})\bar{S}^2}{\sigma^2\gamma} + 1 - 2(s_t - \bar{s})\right] \tag{10.43}$$

Plug in equation (6.14) for $\bar{S}$:

$$r_t^f = -\ln(\delta) + \gamma g - \left(\frac{\gamma}{\bar{S}}\right)^2 \frac{\sigma^2}{2}\left[\frac{2(1-\phi)(s_t-\bar{s})\left[\sigma^2\gamma/(1-\phi)\right]}{\sigma^2\gamma} + 1 - 2(s_t - \bar{s})\right] \tag{10.44}$$

After cancellations:

$$r_t^f = -\ln(\delta) + \gamma g - \left(\frac{\gamma}{\bar{S}}\right)^2 \frac{\sigma^2}{2}\left[2(s_t-\bar{s}) + 1 - 2(s_t - \bar{s})\right] \tag{10.45}$$

$$r_t^f = -\ln(\delta) + \gamma g - \left(\frac{\gamma}{\bar{S}}\right)^2 \frac{\sigma^2}{2} \tag{10.46}$$

Plug in equation (6.14) for $\bar{S}$:

$$r_t^f = -\ln(\delta) + \gamma g - \frac{\gamma^2(1-\phi)}{\sigma^2\gamma}\frac{\sigma^2}{2} \tag{10.47}$$

After cancellations:

$$r_t^f = -\ln\delta + \gamma g - \frac{\gamma}{2}(1-\phi) \tag{10.48}$$

10.9 Appendix I

$$S_{t+1} = e^{s_{t+1}} \tag{10.49}$$

$$C_{t+1} - X_{t+1} = C_{t+1}e^{s_{t+1}} \tag{10.50}$$

$$X_{t+1} = C_{t+1} - C_{t+1}e^{s_{t+1}} = C_{t+1}\left(1 - e^{s_{t+1}}\right) \tag{10.51}$$

$$x_{t+1} = c_{t+1} + \ln\left(1 - e^{s_{t+1}}\right) \tag{10.52}$$

$$\begin{aligned}\frac{dx_{t+1}}{dc_{t+1}} &= 1 + \frac{1}{1 - e^{s_{t+1}}}\left(-e^{s_{t+1}}\lambda\left(s_t\right)\right)\\ &= 1 + \frac{\lambda\left(s_t\right)}{\left(1 - e^{s_{t+1}}\right)\left(-e^{-s_{t+1}}\right)}\\ &= 1 + \frac{\lambda\left(s_t\right)}{-e^{-s_{t+1}} + 1}\\ &= 1 - \frac{\lambda\left(s_t\right)}{e^{-s_{t+1}} - 1}\end{aligned} \tag{10.53}$$

10.10 Appendix J

Using equation (6.18) at $\bar{s}$:

$$1 - \frac{\left(1/\bar{S}\right) - 1}{e^{-\bar{s}} - 1} = 1 - \frac{\left(1/\bar{S}\right) - 1}{\left(1/\bar{S}\right) - 1} = 0 \tag{10.54}$$

10.11 Appendix K

$$1 = E_t\left[M_{t+1}\frac{P_{t+1} + D_{t+1}}{P_t}\right] \tag{10.55}$$

Since price-dividend and price-consumption ratio are equivalent:

$$1 = E_t\left[M_{t+1}\frac{P_{t+1} + C_{t+1}}{P_t}\right] \tag{10.56}$$

$$1 = E_t\left[M_{t+1}\frac{C_{t+1}}{P_t} + M_{t+1}\frac{P_{t+1}}{P_t}\right] \tag{10.57}$$

$$\frac{P_t}{C_t} = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t} + M_{t+1}\frac{P_{t+1}}{C_t}\right] \tag{10.58}$$

$$\frac{P_t}{C_t}(s_t) = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t}\left(1+\frac{P_{t+1}}{C_{t+1}}(s_{t+1})\right)\right]. \tag{10.59}$$

10.12 Appendix L

Let's extend the upper pricing eqution for an asset that lives infinitely long:

$$\frac{P_t}{C_t} = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t}\left(1+\frac{P_{t+1}}{C_{t+1}}\right)\right] \tag{10.60}$$

A two period version is:

$$\frac{P_t}{C_t} = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t}\left(1+E_{t+1}\left[M_{t+2}\frac{C_{t+2}}{C_{t+1}}\left(1+\frac{P_{t+2}}{C_{t+2}}\right)\right]\right)\right] \tag{10.61}$$

Using the law of iterated expectations:

$$\frac{P_t}{C_t} = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t} + M_{t+1}M_{t+2}\frac{C_{t+2}}{C_t} + M_{t+1}M_{t+2}\frac{P_{t+2}}{C_t}\right] \tag{10.62}$$

Now I plug in M_{t+1}:

$$\begin{aligned}\frac{P_t}{C_t} &= E_t\left[\delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma}\frac{D_{t+1}}{D_t}\right]\\ &+ E_t\left[\delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma}\delta\left(\frac{S_{t+2}}{S_{t+1}}\frac{C_{t+2}}{C_{t+1}}\right)^{-\gamma}\frac{C_{t+2}}{C_t}\right]\\ &+ E_t\left[\delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma}\delta\left(\frac{S_{t+2}}{S_{t+1}}\frac{C_{t+2}}{C_{t+1}}\right)^{-\gamma}\frac{P_{t+2}}{C_t}\right]\end{aligned} \tag{10.63}$$

$$\begin{aligned}\frac{P_t}{C_t} &= E_t\left[\delta\left(\frac{S_{t+1}}{S_t}\right)^{-\gamma}\left(\frac{C_{t+1}}{C_t}\right)^{1-\gamma} + \delta^2\left(\frac{S_{t+2}}{S_t}\right)^{-\gamma}\left(\frac{C_{t+2}}{C_t}\right)^{1-\gamma}\right]\\ &+ E_t\left[\delta^2\left(\frac{S_{t+2}}{S_t}\right)^{-\gamma}\left(\frac{C_{t+2}}{C_t}\right)^{-\gamma}\frac{P_{t+2}}{C_t}\right]\end{aligned} \tag{10.64}$$

For infinite time this equations becomes,

$$\frac{P_t}{C_t} = E_t \sum_{j=1}^{\infty} \delta^j \left(\frac{S_{t+j}}{S_t}\right)^{-\gamma} \left(\frac{C_{t+j}}{C_t}\right)^{1-\gamma} \tag{10.65}$$

since $\delta < 1$ will make the last part of the sum go to zero. First we can derive a deterministic version of the equation. Now one plugs in $c_{t+1} - c_t = g$. So $c_{t+j} - c_t = jg$ and $\sigma^2 = 0$. So $s_{t+1} = (1-\phi)\,\bar{s} + \phi s_t$. For j periods:

$$\begin{aligned} s_{t+j} &= (1-\phi)\,\bar{s} + \phi\bar{s} + \phi^j\bar{s} + \phi^j s_t \\ \Leftrightarrow -\left[s_{t+j} - s_t\right] &= -\,(1-\phi)\,\bar{s} - \phi\bar{s} - \phi^j\bar{s} - \phi^j s_t + s_t \\ \Leftrightarrow -\left[s_{t+j} - s_t\right] &= -\,\bar{s} - \phi^j\bar{s} - \phi^j s_t + s_t \\ \Leftrightarrow -\left[s_{t+j} - s_t\right] &= \left(1-\phi^j\right)\left(s_t - \bar{s}\right) \end{aligned} \tag{10.66}$$

In the end the pricing equation becomes:

$$\frac{P_t}{C_t} = \sum_{j=1}^{\infty} \delta^j e^{gj(1-\gamma)} e^{-\gamma(s_{t+j}-s_t)} = \sum_{j=1}^{\infty} \left(\delta G^{1-\gamma}\right)^j e^{\gamma\left(1-\phi^j\right)(s_t-\bar{s})} \tag{10.67}$$

To ensure that the sum does not increase without bound one needs to check that:

$$\lim_{j\to\infty} \left(\delta G^{1-\gamma}\right)^j e^{\gamma\left(1-\phi^j\right)(s_t-\bar{s})} = 0 \quad \forall \quad s_t \in \{0; s_{max}\} \tag{10.68}$$

This condition clearly depends on numerical values for the variables on the left hand side of the upper equation. So I will check that the upper condition holds for all parameter constellations used to calibrate the model. For $\phi < 1$ the second factor is bounded depending on the value of s_t. The first factor has a limit of zero if $\delta G^{1-\gamma} < 1$. So this is the sufficient condition.

10.13 Appendix M

The variable x_t follows a random walk process:

$$x_{t+1} = x_t + e_{t+1}, \quad e_t \sim i.i.d.N\left(0, \sigma^2\right) \tag{10.69}$$

Then follows for x_{t+k}:

$$x_{t+k} = x_t + \sum_{i=1}^{k} e_{t+i} \tag{10.70}$$

The variance scales with time since the shocks are i.i.d.:

$$Var\left[x_{t+k} \mid x_t\right] = Var\left[x_t + \sum_{i=1}^{k} e_{t+i} \middle| x_t\right] \quad \Longleftrightarrow \quad Var\left[x_{t+k} \mid x_t\right] = k\sigma^2 \tag{10.71}$$

and the standard deviation with the square root of time:

$$\sigma\left[x_{t+k} \mid x_t\right] = \sqrt{k}\sigma \tag{10.72}$$

10.14 Appendix N

The variable x_t follows an autoregressive process of order 1:

$$x_{t+1} = \rho x_t + e_{t+1}, \quad e_t \sim i.i.d.N\left(0, \sigma^2\right) \tag{10.73}$$

Then follows for x_{t+k}:

$$x_{t+k} = \rho^k x_t + \sum_{i=1}^{k} \rho^{k-i} e_{t+i} \tag{10.74}$$

The conditional variance of x_{t+k} is:

$$Var\left[x_{t+k} \mid x_t\right] = Var\left[\rho^k x_t + \sum_{i=1}^{k} \rho^{k-i} e_{t+i} \middle| x_t\right] = \sigma^2 \sum_{i=0}^{k} \rho^{2(k-i)} \tag{10.75}$$

if $|\rho| < 1$ then

$$\sigma^2 \sum_{i=0}^{k} \rho^{2(k-i)} < k\sigma^2. \tag{10.76}$$

And so the Sharpe ratio increases faster because the standard deviation appears in the denominator.

List of Figures

List of Tables

Bibliography

[1] Abel, Andrew B.
Risk Premia and Term Premia in General Equilibrium, 1998
Journal of Monetary Economics 43, pp. 3-33

[2] Black, Fischer
Mean Reversion and Consumption Smoothing, 1976
Review of Financial Studies 3, pp. 107-114

[3] Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner
ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, 1992
Journal of Eonometrics 52, pp. 5-59

[4] Campbell, John Y. and Robert J. Shiller
The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, 1988a
Review of Financial Studies 1, pp. 195-227

[5] Campbell, John Y. and Robert J. Shiller
Stock Prices, Earnings, and Expected Dividends, 1988b
Journal of Finance 43, pp. 661-676

[6] Campell, John Y. and John H. Cochrane,
Explaining the Poor Performance of Consumption-Based Asset Pricing Models, 1998a
unpublished paper, Harvard University and University of Chicago

[7] Campell, John Y. and John H. Cochrane,
Appendix to 'By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, 1998b
unpublished paper, Harvard University and University of Chicago

[8] Campbell, John Y. and John H. Cochrane,
By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, 1999
Journal of Political Economy 107, pp. 205-251

[9] Campbell, John Y. and John H. Cochrane,
By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, 1995
NBER Working Paper No. 4995

[10] Campbell, John Y.
Asset Prices, Consumption, and the Business Cycle, 1998
Handbook of Macroeconomics, North-Holland: Amsterdam

[11] Chapman, David A.
Habit Formation and Aggregate Consumption, 1997
Econometrica 66(5), pp. 1223-1230

[12] Cochrane, John H.,
Asset Pricing, 2001
Princeton University Press: Princeton

[13] Cochrane, John H. and Lars P. Hansen,
Asset Pricing Lessons for Macroeconomics, 1992
in Olivier Blanchard and Stanley Fischer, Eds. 1992 NBER Macroeconomics Annual

[14] Fama, Eugene F. and Kenneth R. French,
Business Conditions and Expected Returns on Stocks and Bonds, 1989
Journal of Financial Economics 25, pp. 23-49

[15] Fama, Eugene F. and Kenneth R. French,
Permanent and Temporary Components of Stock Prices, 1988a
Journal of Political Economy 96, pp. 246-273

[16] Fama, Eugene F.
Stock Retruns, Expected Returns, and Real Activity, 1990
Journal of Finance 45, pp. 1089-1108

[17] Ferson, Wayne E. and George Constantinides
Habit Persistence and Durability in Aggregate Consumption: Empirical Tests, 1991
Journal of Financial Economics 29, pp. 199-240

[18] Gourieroux, Christian and Joann Jasiak
Financial Econometrics, 2001
Princeton University Press: Princeton

[19] Heaton, John C.
An Empirical Investigation of Asset Pricing with Temporary Dependent Preference Specifications, 1995
Econometrica 63, pp. 681-717

[20] Mas-Colell, Andreu, Michael D. Whinston, Jerry R. Green,
Microeconomic Theory, 1995
Oxford University Press: Oxford

[21] Mehra, Rajnish and Edward C. Prescott,
The Equity Premium: A Puzzle, 1985
Journal of Monetary Economics 15, pp. 145-161

[22] Meyer, Bernd
Intertemporal Asset Pricing: Evidence from Germany, 1999
Physika-Verlag: Heidelberg

[23] Miranda, Mario J., Paul L. Fackler,
Applied Computational Economics and Finance, 2002
The MIT Press: Cambridge, MA

[24] Sundaresan, Suresh M.
Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth, 1989
Review of Financial Studies 2, pp. 73-88